[redacted]

ELSEVIER
PERGAMON

ENCYCLOPEDIA OF LANGUAGE & LINGUISTICS

SECOND EDITION

EDITOR-IN-CHIEF
KEITH BROWN

COORDINATING EDITORS
ANNE H. ANDERSON
LAURIE BAUER
MARGIE BERNS
GRAEME HIRST
JIM MILLER

ELL2

Volume 6

Inupiaq -
Learning Sign Language

Inupiaq

L D Kaplan, University of Alaska, Fairbanks, AK, USA

Introduction and Dialectology

Inupiaq (Inupiatun, North Alaskan and Inupiatun, Northwest Alaska) is an Eskimo-Aleut language spoken in Alaska, part of the Inuit branch or Eastern group of Eskimo languages and distinct from the Yupik languages. Extending across the Arctic from Alaska through Canada to Greenland, Inuit varieties differ from each other in significant ways but nowhere is there found a sharp internal break that would constitute a language border, and so they are considered a dialect continuum (Figure 1). The Inuit groups of the Eastern Arctic are a diaspora that spread from Alaska, and the Bering Sea area is recognized as the homeland of the Eskimo-Aleut language family and people (*see* **Eskimo-Aleut**).

Inupiaq comprises two dialect groups, North Alaskan Inupiaq (NAI) and Seward Peninsula Inupiaq (SPI), each with two dialects. NAI includes the North Slope and the Malimiut dialects, with the former spoken along the Arctic coast and the latter found around Kotzebue Sound, along the Kobuk River and south at the head of Norton Sound. SPI includes the Qawiaraq dialect, found along the northern shore of Norton Sound, on the southeastern Seward Peninsula and at Teller, and the Bering Strait dialect is found along the shores of Bering Strait and on the offshore islands, King Island (now uninhabited), and Little Diomede. Dialects are distinguished primarily in terms of phonology, lexicon, and morphology and include a number of subdialects.

Phonology and Writing

The Alaskan Inupiaq writing system was developed by Roy Ahmaogak, a Barrow Inupiaq minister, and linguist Eugene Nida in 1946 and has undergone revisions since. Current orthographic symbols are given below with equivalent phonetic symbols in parentheses where the two differ. The entire palatal series is absent in SPI, *z* is present only in the Bering Strait dialect, and *e* has limited occurrence, found in SPI only, particularly on Little Diomede Island.

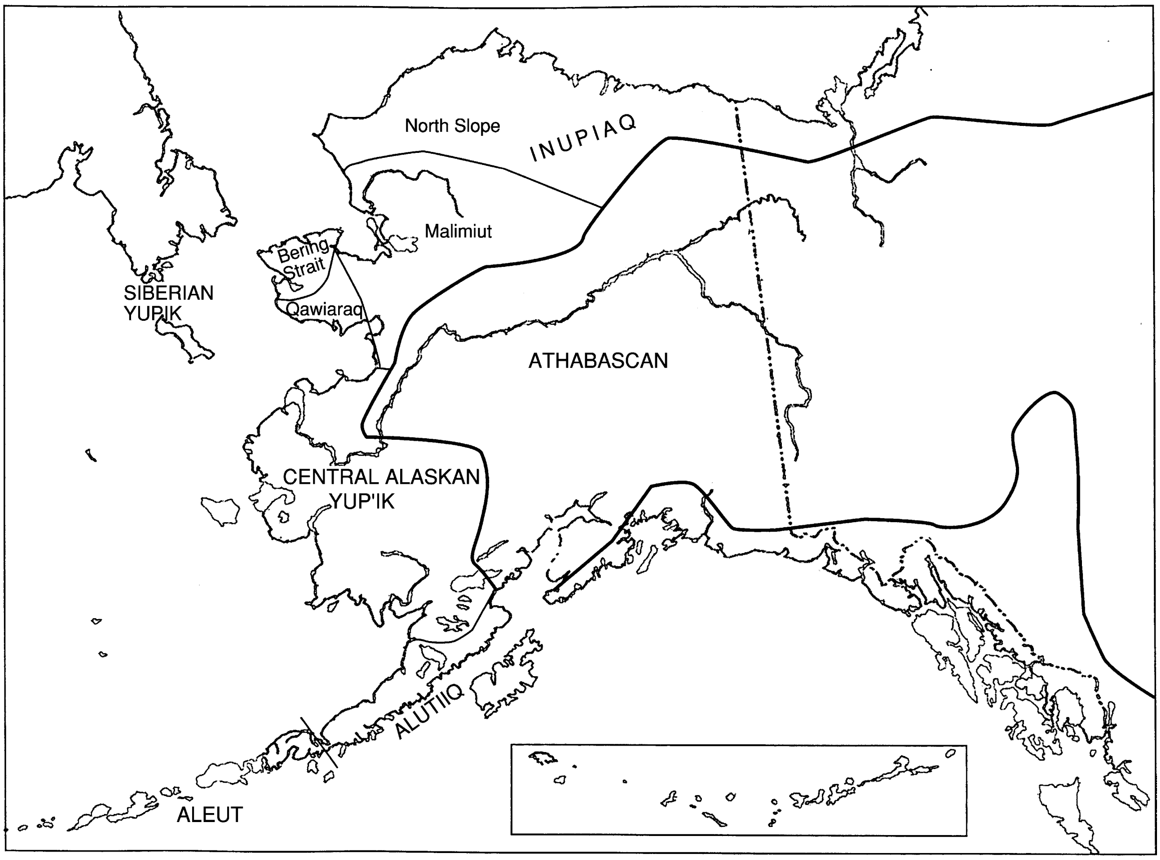


Figure 1 Eskimo-Aleut languages of Alaska.

Consonants:

p	t	ch(č)	k	q	‘(ʔ)
	ł(ɬ̥)	ł̣(ɭ̥)	ḷ(lʲ), y	g (ɣ)	ġ(ʀ)
m	n	ñ	ŋ		

Vowels and diphthongs:

i	e(ɨ)	u
		a

Consonants and vowels can be long and are then written double. Any two vowels with the exception of *e* may be paired: *ai ia ua au iu ui*. In most of Malimiut, certain diphthong pairs have coalesced and are pronounced identically: *ia* and *ai* are pronounced [e:], *au* and *ua* are pronounced [o:], and *iu* is pronounced like *ii* as [i:]. A major phonological phenomenon in Inuit languages is consonant assimilation, which increases in magnitude as one travels east (*see* **West Greenlandic**). Alaskan Inupiaq has differentiated clusters, although the North Slope dialect has assimilation by manner of articulation. Palatalization of alveolar consonants is notable in NAI, with palatalization of some velars found in much of Malimiut. A major distinguishing feature of the Inuit branch is the simplification of the Proto-Eskimo four-vowel system by the merger of the central vowel ɨ, principally with i. Reflexes of historical *i* trigger progressive palatalization of alveolars and assibilation of prevocalic *t* in NAI: NAI *iñuk* ‘person’ but SPI *inuk* and NAI *isiq-* ‘enter’ but SPI *itiq-*. Reflexes of the fourth vowel ɨ are typically [*i*] but may undergo deletion or alternation with other vowels.

Morphology and Syntax

Inupiaq is a highly polysynthetic language, and suffixation creates very long words, so that a word is often equivalent to an entire English sentence. Words are typically constructed of a noun or verb stem, one or more derivational suffixes, and an inflectional ending, and may be singular, dual, or plural in number. Examples are from the North Slope dialect:

(1) atuġ-nia-ŋit-chuk
sing-FUT-NEG-3du/INTRANS
‘they (two) will not sing’

Verbs may be intransitive or transitive, and transitive verb endings mark number and person of both subject and object.

(2) tautuk-kaat
see-3plSUBJ/3singOBJ
‘they see him/her/it’

The verb ending here expresses the subject and direct object, and pronouns are unnecessary, although they can be added in, as can nouns. Pronouns express person and number but not gender, which is expressed only by nouns. There are no articles.

(3) aŋutit aġnaq tautuk-kaat
men.pl woman see-3plSUBJ/3singOBJ
‘the men see the woman’

Word order is Subject-Object-Verb, although the complex inflectional system makes free word order possible. Through the influence of English, many speakers now prefer SVO order.

Lexicon and Ethnonyms

Lexical borrowing is primarily from Russian and English. Although the Russians never had a permanent presence in Inupiaq Alaska, the dozen or so Russian borrowings found in Inupiaq were probably introduced through trade with mainland Yupiks to the south and are found principally in SPI and Malimiut. Early English borrowings were introduced beginning in the late 19th century through contact with traders and whalers and were well integrated phonologically (e.g., *palauvak* ‘flour’). Modern-day bilingualism often gives rise to use of English words in Inupiaq speech, but these occurrences cannot be considered true borrowings.

Besides indicating the name of the language, ‘Inupiaq’ can also be used in English either as an adjective (e.g., the Inupiaq language) or a singular noun for a person (e.g., an Inupiaq from Barrow). The plural ‘Inupiat’ also is used for people (the Inupiat people, the North Slope Inupiat). With consonant palatalization, NAI uses the spellings ‘Iñupiaq’ and ‘Iñupiat,’ whereas SPI lacks palatalization and uses ‘Inupiaq’ and ‘Inupiat.’ Alaskans also use ‘Eskimo,’ although this term is disfavored in Canada.

Population and Viability

There are some 13 500 Inupiat in Alaska, about 3000 of whom speak the Native language. Inupiat are now bilingual or speak only English. Most speakers are in their late forties or older, and in some areas the youngest Inupiaq speakers are in their sixties or even seventies. The language shift to English is brought about by a number of factors: government, education, and media are largely in English (although Inupiaq is often heard on the radio); monolingual English speakers have lived among the Inupiat for decades; and airplanes have made travel outside the area easy. In addition, past Inupiaq language use was discouraged and often punished by teachers and school officials. In 1998, a majority voted to make English Alaska’s official language, sending a negative message in the view of Native language supporters.

The continued use of Inupiaq as a spoken language is threatened, and efforts at revitalization consist largely of school language classes. Language immersion programs exist in elementary schools in Kotzebue and Barrow, from which it is hoped that new generations of Inupiaq speakers will emerge.

See also: Canada: Language Situation; Eskimo-Aleut; West Greenlandic; United States of America: Language Situation.

Bibliography

Berge A & Kaplan L (in press). 'Contact-induced lexical development in Yupik and Inuit languages.' *Etudes Inuit Studies*.

Fortescue M, Jacobson S & Kaplan L (1994). *Comparative Eskimo dictionary with Aleut cognates*. Fairbanks: Alaska Native Language Center.

Kaplan L D (1981). *Phonological issues in North Alaskan Iñupiaq*. Research paper no. 7. Fairbanks: Alaska Native Language Center.

MacLean E (1980). *Abridged Iñupiaq and English dictionary.* Fairbanks: Alaska Native Language Center and Barrow: Inupiat Language Commission.

MacLean E (1986a). *North Slope Iñupiaq Grammar: first year.* Fairbanks: Alaska Native Language Center.

MacLean E (1986b). *North Slope Iñupiaq Grammar second year.* Fairbanks: Alaska Native Language Center.

Webster D H & Zibell W (1970). *Iñupiat Eskimo Dictionary.* Fairbanks: Summer Institute of Linguistics.

Invented Language in Literature

P Stockwell, University of Nottingham, Nottingham, UK

Invented Language and Literary Worlds

All creative language, articulated in literary texts, is invented language, of course. However, all creative fiction depends on a greater or lesser degree of difference between our actual world and the imagined world of the text. Even a highly realist and naturalistic narrative, even a documentary faction or even a historical reconstruction are all likely to depart in small ways from the hard reality, and the form of expression used to articulate the divergent world will need to be different (and invented) from the language that is **really** real. As soon as a proper name for a character ('Snoopy,' 'Hamlet,' 'Heathcliff') is used, this invention serves to mark the divergence of worlds from that of the reader. Even words with apparently stable referents are used inventively in literary fiction: Dicken's 'London' is not our London; Philip Pullman's different 'Oxfords' are explicitly not our Oxford; even Wilfred Owen's Great War trenches are his literary projection of that landscape, rather than the landscape itself.

Furthermore, literary authors tend to express themselves in ways that are commonly regarded as 'poetic' or 'expressive.' This stylistic creativity can be manifest as metaphorical compounds (the Old English 'mere-stape' – 'sea-stepper' for a ship), or as lexical blends (e.g. cummings's 'manunkind'), or as grammatical disjunction (Robert Heinlein's 'the door dilated'), and so on. In most literature, though, these examples of departures from everyday language are minimal and serve only to mark out a local effect. They are usually one-off creations rather than being part of a thoroughly worked-out and systematic linguistic structure.

In this article, more radical linguistic inventiveness will be discussed, ranging from the creation of sets of new words that have a strong iconic value, to the creation of fully-fledged invented language systems. The functions of such creative artifacts will also be discussed, but it is worth pointing out at the beginning that the primary function of new language in literary fiction is to delineate the distance and connections between the reader's world and the world imagined in the text. Whether the invented language is for local ornamental color or seems intended for a central thematic purpose, imaginary lexis, grammar, and pragmatic patterns play an important role in establishing the readerly perception of escapism or political comment. Delineating the difference between the textual world and the reader's world is the first step in determining the significance of the work, in making identifications for empathy or satire, in being able to generalize the specifics of the storyworld onto principles for the reader's world.

This article begins with the smallest scale of linguistic inventiveness, in new words and new uses for existing words, before going on to consider higher level discourse creativity, and then full invented language systems. In setting out this scale, however, it is important to realize that a few well-placed new words can have as significant an effect as thoroughly imagined rich linguistic innovation. Therefore, this article will consider the artistic function of inventiveness separately. The article ends with a consideration of

The continued use of Inupiaq as a spoken language is threatened, and efforts at revitalization consist largely of school language classes. Language immersion programs exist in elementary schools in Kotzebue and Barrow, from which it is hoped that new generations of Inupiaq speakers will emerge.

See also: Canada: Language Situation; Eskimo-Aleut; West Greenlandic; United States of America: Language Situation.

Bibliography

Berge A & Kaplan L (in press). 'Contact-induced lexical development in Yupik and Inuit languages.' *Etudes Inuit Studies*.

Fortescue M, Jacobson S & Kaplan L (1994). *Comparative Eskimo dictionary with Aleut cognates.* Fairbanks: Alaska Native Language Center.

Kaplan L D (1981). *Phonological issues in North Alaskan Iñupiaq*. Research paper no. 7. Fairbanks: Alaska Native Language Center.

MacLean E (1980). *Abridged Iñupiaq and English dictionary.* Fairbanks: Alaska Native Language Center and Barrow: Inupiat Language Commission.

MacLean E (1986a). *North Slope Iñupiaq Grammar: first year.* Fairbanks: Alaska Native Language Center.

MacLean E (1986b). *North Slope Iñupiaq Grammar second year.* Fairbanks: Alaska Native Language Center.

Webster D H & Zibell W (1970). *Iñupiat Eskimo Dictionary.* Fairbanks: Summer Institute of Linguistics.

Invented Language in Literature

P Stockwell, University of Nottingham, Nottingham, UK

Invented Language and Literary Worlds

All creative language, articulated in literary texts, is invented language, of course. However, all creative fiction depends on a greater or lesser degree of difference between our actual world and the imagined world of the text. Even a highly realist and naturalistic narrative, even a documentary faction or even a historical reconstruction are all likely to depart in small ways from the hard reality, and the form of expression used to articulate the divergent world will need to be different (and invented) from the language that is **really** real. As soon as a proper name for a character ('Snoopy,' 'Hamlet,' 'Heathcliff') is used, this invention serves to mark the divergence of worlds from that of the reader. Even words with apparently stable referents are used inventively in literary fiction: Dicken's 'London' is not our London; Philip Pullman's different 'Oxfords' are explicitly not our Oxford; even Wilfred Owen's Great War trenches are his literary projection of that landscape, rather than the landscape itself.

Furthermore, literary authors tend to express themselves in ways that are commonly regarded as 'poetic' or 'expressive.' This stylistic creativity can be manifest as metaphorical compounds (the Old English 'mere-stape' – 'sea-stepper' for a ship), or as lexical blends (e.g. cummings's 'manunkind'), or as grammatical disjunction (Robert Heinlein's 'the door dilated'), and so on. In most literature, though, these examples of departures from everyday language are minimal and serve only to mark out a local effect. They are usually one-off creations rather than being part of a thoroughly worked-out and systematic linguistic structure.

In this article, more radical linguistic inventiveness will be discussed, ranging from the creation of sets of new words that have a strong iconic value, to the creation of fully-fledged invented language systems. The functions of such creative artifacts will also be discussed, but it is worth pointing out at the beginning that the primary function of new language in literary fiction is to delineate the distance and connections between the reader's world and the world imagined in the text. Whether the invented language is for local ornamental color or seems intended for a central thematic purpose, imaginary lexis, grammar, and pragmatic patterns play an important role in establishing the readerly perception of escapism or political comment. Delineating the difference between the textual world and the reader's world is the first step in determining the significance of the work, in making identifications for empathy or satire, in being able to generalize the specifics of the storyworld onto principles for the reader's world.

This article begins with the smallest scale of linguistic inventiveness, in new words and new uses for existing words, before going on to consider higher level discourse creativity, and then full invented language systems. In setting out this scale, however, it is important to realize that a few well-placed new words can have as significant an effect as thoroughly imagined rich linguistic innovation. Therefore, this article will consider the artistic function of inventiveness separately. The article ends with a consideration of

the linguistic assumptions that seem to underlie these linguistic experiments. Since most examples of invented language come from the genres of science fiction and fantasy, the way that sf authors have imagined the science of linguistics is relevant to the discussion.

Neologism and Neosemy

New words ('neologisms') are a particular feature of science fiction (sf) and fantasy literature, where they have also been termed 'lexoids' (Mary Ellen Ryder, personal communication). Sf filled its early adventure stories with 'ray-guns,' 'ion-drives,' and 'teleporters,' and featured 'warp-speed' 'starships,' 'flying saucers,' and 'cyborgs.' There are clearly several types of word-formation at work here, drawing on all of the available resources of lexical generation in English. The following typology of neologisms is drawn from Stockwell (2000: 115–138).

Perhaps the most common form (unsurprisingly for a Germanic-based language like English) is compounding. Here, root words can be placed together as a single word lexical blend ('teleport,' 'holodeck,' 'cyberspace'). The most common source for such compounds draws on noun-combinations (N+N), as in 'wormhole' and 'starship' and hyphenated forms, such as 'ion-gun' and 'rim-world.' Because of the commonality of the source lexis here, it is fairly easy to work out a meaning for such neologisms, especially as they are always used in a disambiguating narrative context. N+V forms ('warp-drive,' 'terraform') also occur, and even Adj+V ('Newspeak') and V+Adj ('spindizzy'). Single-word neologisms such as these are typical of technologically-based science fiction, where a new invention or technical phenomenon requires an invented and brief label.

Compounding can also take the form of multi-word lexical items, such as 'flying saucer,' 'dilithium crystal,' 'tractor beam,' 'force field,' and many others. These, too, are very common in science fiction as they carry the flavor of a technical specification as well as often drawing on a scientific register. A less common form of phrasal circumlocution is also found, where the sense of 'newness' is diminished because the source elements are easily recognized. Examples such as 'sense-of-we-feeling,' 'Bug Eyed Monster,' or 'cultural engineering' retain a sense of a familiar phenomenon modified in the imagined world. All of these examples are fairly easy to understand.

Neologisms can also be created by borrowing. In this case, the borrowing can be from either terrestrial or imagined alien languages – though of course often the imagined alien language bears uncanny resemblance to root morphemes in terrestrial languages. In early sf, words were often borrowed from languages perceived by the British or American writers to be exotic. The proto-science fictional *Historie of the man in the moone* (1634) by Bishop Francis Godwin has the moon people speaking 'Lunarian,' based on his rather sketchy understanding of how Chinese sounded.

Alien borrowing (the 'Daleks' from *Dr Who,* or the 'Wookies' from *Star Wars*) can always be imagined as having terrestrial influences, of course. By far the most common source languages for neologistic borrowing are Greek and Latin, because of their use in much technical terminology and perhaps also because the decline in classical education over the last century has rendered the etymology more obscure and exotic. 'Sol,' 'Luna,' and 'Terra' have become common names for our local star system. 'Android,' 'cybernetics,' and 'cryogenics' all have Greek origins.

Straightforward original creation is rare, since the etymological origins of neologisms can usually be perceived. Usually only proper names show real creation, and these are then turned into full neologisms by grammatical derivation: 'Odonism,' 'Naderites,' 'Morlocks.' Such derivation also leads to 'back-formation,' where analogous forms are created: 'white holes,' 'corpsicle,' and 'astrogation.' The last two of these also illustrate longer coinages being abbreviated – it is common to find such shortened back-formations breaking the usual morphological boundaries: 'droid' is a broken form of 'android'; and the malicious race of 'Borg' are broken down from 'cyborgs,' which is itself an abbreviated and blended version of 'cybernetic organism,' which in turn consists of two words drawn from Greek.

Shortened forms like these, together with abbreviations, are another means of creating neologisms. 'Fax' (from 'facsimile'), 'bionics' (from 'biological electronics'), and 'cyborg' are all shortened forms that have become perceived as lexical items in their own right. Acronyms formed from the initial letters of phrases have also become lexicalized, and pronounced either 'atomically' in their separate elements ('UFO,' 'ESP,' 'ET') or 'molecularly' as a phonetic sequence (as in 'ufologist,' one who investigates unidentified flying objects; and 'esper' who has the power of extra-sensory perception; and the research organization 'SETI,' searching for extraterrestrial intelligence).

Of course, all the usual morphological inflections of English (such as the 's'-plural, the 'ing'-participle, the 'ed'-past suffix, and so on) further extend most of these neologisms. It is rare, however, with single neologisms to find inflectional creativity in literature. There is no parallel of the historical shift in

English from the third-person verb suffix '-eth' to '-s' ('maketh' to 'makes'), for example. Where there is innovation in grammatical particles, these tend to form part of a fully-fledged imagined linguistic system, rather than simply being associated with isolated neologisms.

Aside from neologisms proper, literary fiction tends to rely on 'neosemy' (a term coined in Stockwell, 2000), in the sense of a new meaning developed for an existing word. The neologism 'utopia' for example (from Thomas More's 1516 work (1910)) was coined as a proper name for his idealized island but has since been extended to refer to any idealized state, and inflected into 'utopian,' and back-formed into 'dystopian' and 'heterotopian.' A 'flying saucer' and a 'black hole,' strictly speaking, should be regarded as neosemes rather than neologisms. Neosemes have the advantage in literary fiction of making the imagined world seem close and more nearly extrapolated from the reader's familiar world. Neologisms, by contrast, can seem relatively disjunctive and alien.

Dialectal Extrapolation

More extensive than neologistic innovation but falling short of a full invented language, dialectal invention is a common compromise in literary fiction. Imagining the appropriate register or dialect that applies in the projected world allows the texture of difference to be established while not making undue demands on the reader to engage in detailed decoding. Sometimes, the dialectal projection is simply a product of a high density of neologisms and neosemes, drawn from source lexis that is reasonably easy to recognize. The futuristic dialect of William Gibson's *Sprawl* series of novels works like this (in fact, it was developed largely from 1970s Canadian biker slang):

> Molly was snoring on the temperfoam. A transparent cast ran from her knee to a few millimeters below her crotch, the skin beneath the rigid micropore mottled with bruises, the black shading into ugly yellow. Eight derms, each a different size and colour, ran in a neat line down her left wrist. An Akai transdermal unit lay beside her, its fine red leads connected to input trodes under the cast.
>
> He turned on the tensor beside the Hosaka. The crisp circle of light fell directly on the Flatline's construct. He slotted some ice, connected the construct, and jacked in. (Gibson, 1984: 98–99)

This passage, from the novel *Neuromancer* (itself a punning lexical blend neologism), displays neologisms that are easy to decode in isolation ('micropore'), in context ('snoring on the temperfoam'), and with a little etymological knowledge ('derm'). There are abbreviations that would be familiar to sf readers ('trodes' from 'electrodes'), as well as some neosemic usage: 'ice' as access to cyberspace, 'construct' as a noun for a software program, and 'jacked' applied to a human body rather than a machine's cable.

Similarly, Jeff Noon's 1993 novel *Vurt* draws on the street-slang of early 1990s Manchester, with some neologistic blends:

> Shadowcop! Broadcasting from the store wall, working his mechanisms; flickering lights in smoke. And then a flash of orange; an inpho beam shining out from the shadowcop's eyes. It caught Mandy in its flare-path, gathering knowledge. She ducked down from the beam, banging, hard-core, on the van doors. (Noon, 1993: 7)

Here the new phenomena are explained by the context, and rather than a clumsy definition being supplied, information is presented appositionally ('gathering knowledge') that is consistent with the point of view of the narrator. Both of these examples of dialectal extrapolation serve not only to gesture towards the alternate world, but also serve as part of the characterization in the texts. It is noticeable that literary fictions that feature dialectal innovation tend to be strongly focalized through a participating character: the narrative is thus their dialect, and the invented language takes the reader quickly into the conscious world of the novel.

Such dialectal invention is fairly minimal, of course, and some writers have developed the principle of dialectal extrapolation to underline the alternativity of their imagined worlds. Anthony Burgess, in his 1962 novel *A clockwork orange*, creates a dialectal slang for his violent gang-characters Alex and his 'droogs' (friends). Like many lexical items in the novel, this word derives from transliterated or anglicized Russian, along with 'baboochka' (old woman), 'bezoomny' (mad), 'bitva' (battle), 'gulliver' (head), 'horrorshow' (good), 'koshka' (cat), 'krovvy' (blood), 'malchick' (boy), 'skazzed' (said), 'voloss' (hair), and many others. The slang is called 'Nadsat,' the suffix equivalent to 'teen' at the end of the Russian words for the numbers 11 to 19. Nadsat also includes some Cockney-derived words and other inventions: 'viddy' (see), 'pretty polly' (money), 'platties' (clothes), and so on.

Nadsat serves to increase the reader's involvement with the focalizer, Alex, but replacing many of the words also allows scenes of 'ultra-violence' and rape to be portrayed with an immediacy that the reader might otherwise recoil from more instantly. In this way, the novel pitches the reader into Alex's mind

more effectively than if the narrative had been written entirely in Standard English.

The dialectal style serves a thematic purpose in the novel, rather than being merely ornamental or part of the background enrichment of the imagined world. Similarly, Russell Hoban's *Riddley Walker* (1980) presents an even more thoroughly realized dialectal extrapolation to symbolize the postapocalyptic fall of civilization, as seen through the eyes of the eponymous protagonist:

> Stoans want to be lissent to. Them big brown stoans in the formers feal they want to stan up and talk like men. Some times youwl see them lying on the groun with ther humps and hollers theywl say to you, Sit a wyl and res easy why dont you. Then when youre sitting on them theywl talk and theywl tel if you lissen. Theywl tel whats in them but you wont hear nothing what theyre saying without you go as fas as the stoan. You myt think a stoan is slow thats becaws you wont see it moving. (Hoban, 1980: 158)

Here, the written language of the future is imagined as a return to spoken vernacular forms with grammatical patterns reminiscent of existing rural nonstandard southern British English (the demonstrative 'them,' negative concord 'wont hear nothing,' relativizer 'what theyre saying'). The effect is to convey a simple-minded consciousness through which the reader must reconstruct the historical catastrophe that has befallen the country. The extrapolation relies on a set of assumptions about language change, of course, and an identification of spoken dialect with simplicity and social regression. Such extrapolation is artistically successful even if it is sociolinguistically dubious.

Hoban's style requires a great deal of effort on the part of the reader, with a corresponding risk that the reader might consider the task too onerous. The tactic has been pushed to its extreme by Brian Aldiss, in his psychedelic novel *Barefoot in the head* (1969). Here, he imagines a Europe bombed by psychotropic drugs that gradually disorientate the main character, Charteris. As the novel progresses, the lexicogrammatical patterning gradually comes to mirror Charteris's hallucinogenic perception:

> In the square they milled and sledded letting their origins down with mood music thrombic. The body hair buttressed and limbs rebuddied. Metamorphin slipper waker-slip. What they heard they herded and sluice-jiuces ran underfoot. As he luted their animinds. Ages went down into oceanic undertow. Civilians poured in and the old grey and biscuit buildings titaniced down into a glacial cobblesea. Inunvation of hands, plattening feet of limbic brand. Churning flowermotion with eddying scuffles sob-streperous among the one-backed beats. (Aldiss, 1969: 210)

More than 200 pages of this prose are incredibly difficult to read, but the effect can be rewarding if a great deal of attentive effort is expended. The transformations rely largely on lexical blending and collocational clashes, but for the most part the language is still recognizably derived from English. However, further linguistic experimentation has moved beyond existing natural languages altogether, an area that this article will consider in the next section.

Neographies

Inventing new languages is not simply the preserve of literary fiction. The Internet features many hundreds of 'neographies' that have been constructed either as part of literary or cinematic worlds, as extensions of such languages as imagined by fans of the originals, as experiments for teaching aspects of linguistics, or simply as the hobby of numerous people who enjoy inventing new principled artificial language systems. (Enter 'conlang' as an Internet search term to find most of these). By contrast with natural languages, artificial or constructed languages have no real native speakers – though some, like Ludovic Zamenhoff and Clara Zilbernik's Esperanto and its predecessor Johann Schleyer's Volapük, aspire to be used as genuine international languages.

Planned languages or constructed languages ('conlangs') are regarded by their enthusiasts as falling into three categories. Firstly, auxiliary languages ('auxlangs') are those like Esperanto or stripped down technical languages like Basic English, Business English, or Air Traffic Control English. Secondly, languages developed for artistic purposes ('artlangs') encompass invented languages that appear in literary works, as well as invented languages that appear in their own supporting worlds, and invented languages existing only for their own sake. The last two types are particularly prevalent as pastimes posted on the Internet. Thirdly, logical languages ('loglangs') are those devised usually by linguists, philosophers, or cultural critics as a means of demonstrating or teaching some point of principle in linguistics or philosophy – though most computer program languages would also count as loglangs.

For example, from 1955 the sociologist and science fiction writer James Cooke Brown invented a loglang called 'Loglan,' created to test out the Sapir-Whorf hypothesis in linguistics (which, briefly, draws a strong connection between linguistic structure and the patterns of thought that are then determined and constrained by language). The aim was to be as regular, principled, fixed, and stable as possible, since failures in these features in natural languages, it was argued, were what led to ambiguity and logical

lapses. If the strong form of the Sapir-Whorf hypothesis were true, then the speaker of Loglan should be a more logical thinker than the speaker of a natural language. The Logical Language Group released a public-domain version of the language as Lojban in 1987. Loglan and Lojban are based on predicate logic, with predicate functions (roughly, verbs) and participants (roughly, nouns); there are no adjectives or adverbs. Lojban has a greater vocabulary, with the aim of becoming a widely-used expressive language.

The science fiction writer Robert Heinlein was attracted to logical languages such as Loglan (and in fact refers directly to it in his 1966 novel *The moon is a harsh mistress*). In his 1949 novella, 'Gulf,' he also imagines a language called 'Speedtalk,' in which whole basic English words are represented by single syllables or phonemes. This brevity cuts down on the length of utterances required to set out a proposition and allows faster communication of complex ideas. In the story, the language develops superintelligence. Heinlein does not give extended examples of Speedtalk, of course; his artistic purpose is satisfied merely by gesturing towards the existence of the language rather than by realizing it fully. Loglan, by contrast, aims at full transparency and manifestness. This distinction – based on how fully-developed the language is – seems a more useful scale than the loglang/artlang distinction, based on aims and purposes that cannot always be known.

'Neographies,' the common Internet term for invented languages, points to the written nature of most constructed languages. As in most education systems in the world, principles that seem to apply to written language are applied exactly to spoken discourse, and literary authors imagine their languages – even when presenting speech forms – in terms appropriate to writing. This style is not surprising given that the languages involved are expressed in literary text and have no real native speakers to enunciate them. However, it also foregrounds the symbolic or emblematic value of the language itself rather than its value simply as a medium of communication, which spoken discourse foregrounds.

In George Orwell's (1948) *1984*, for example, the main character is employed on the 'Newspeak' dictionary. Newspeak is the version of English aimed at by the repressive state apparatus of IngSoc and Big Brother in the novel: its objective is to cut down the vocabulary of the language in order to diminish the capacity for complex or subversive thought. Clearly, the principles of Newspeak fly in the face of much current linguistic orthodoxy. The idea ignores the ever-shifting processes of linguistic change, ignores dialectal variation, regards dictionaries as comprising and fixing the lexicon, and makes a strong Sapir-Whorfian connection between language and the determination of thought. It also deals only with vocabulary and does not say much about syntactic expression, and it does not deal with the subversive and creative possibilities of metaphorical extension and compounding.

Of course, none of this matters, because Orwell was not extrapolating the future but caricaturing his present in 1948. Newspeak stands as a warning about linguistic impoverishment and the dangers of rhetorical manipulation by politicians and institutions. The novel itself is not written in Newspeak; presumably its political message could not be rendered in the language. No character in the novel even expresses themselves in an extended way in Newspeak: most of the material on the language is given in an appendix to the novel, and only the odd few words ('doublethink,' 'plusgood,' 'duckspeak') are used in the course of the narrative itself. Newspeak is a neograph that is gestured towards for its symbolic value rather than enacted.

Where invented languages are mentioned but do not appear fully, it is usually because they are imagined to be so far in advance of the reader's pitiful mind that the narrator cannot render them in a way which would be at all comprehensible. In Greg Bear's *Thistledown* series, for example, beginning with the 1985 novel *Eon*, the humans from the distant future have dispensed with the monodimensionality of natural language and augmented it with visual projected images and symbols:

> The President's advocate was not pleased with the news Olmy brought.
>
> 'The P.M. had no suspicion of this when you alone were sent?' Toller picted. The symbols that flashed between the two men came from pictor torques around their necks, devices which generated and projected the graphicspeak that had developed over the centuries in the Thistledown and in the Axis City. (Bear, 1985: 258)

'Picting' can only be referred to, not shown, in this way. The implication is that picting is not only far superior in propositional efficiency but also allows a greater precision of complex emotions and shading to be applied to communication.

This indirect gesturing is common for invented languages, with Thomas More adopting the tactic in his 1616 *Utopia*. At the end of the text, he adds "A metre of four verses in the Utopian tongue":

> Utopos ha Boccas peu la chama polta chamaan.
> Bargol he maglomi baccan soma gymnosophaon.
> Agrama gymnosophon labarem bacha bodamilomin.
> Volvala barchin heman la lavolvala dramme pagloni.

Looking suspiciously like a blend of morphemes from modern European languages, particularly Italian and Greek, these phrases are 'translated' into an English verse that takes up twice as many lines. Clearly the primary point is to emphasize the economy of the Utopian language and the poverty in which it can be only "rudely englished." The text ends with a note from "The printer to the reader," in which he apologizes for not reproducing the characters of the Utopian alphabet since, strange as they are, he does not have them yet.

By contrast, in his *Culture* novels, Iain M. Banks provides the graphology for his synthetic galactic language, Marain, but adopts the same strategy of withholding the lexicogrammar as being too advanced for contemporary readers:

> Those of you unfortunate enough not to be reading or hearing this in Marain may well be using a language without the requisite number or type of personal pronouns, so I'd better explain that bit of the translation.
>
> Marain, the Culture's quintessentially wonderful language (so the Culture will tell you), has, as any schoolkid knows, one personal pronoun to cover females, males, in-betweens, neuters, children, drones, Minds, other sentient machines, and every life-form capable of scraping together anything remotely resembling a nervous system and the rudiments of language. (Banks, 1988: 99)

Here, he gestures towards the language without actually revealing it. However, at the end of the novel, he does sign off the narrator using Marain graphology, the symbols of which comprise angular lines on a square grid and look to a nonreader like Japanese katakana graphemes. The alien writing is followed by the name of the narrator, and it is easy to work out that there is a phonetic correspondence between the symbol and the name.

In a note posted to several Internet newsgroups and which will appear in the second edition of his 1991 short story collection, *The state of the art*, Banks has worked out this systematic graphological system. A nine-bit binary code such as 'xooxxxxoo' can be drawn as a 3 x 3 grid:

xoo
xxx
xoo

This graph then resembles the Marain grapheme that looks like an alphabetical 'T' on its side – this represents the phoneme /w/. There are 512 possible symbols on the 3 x 3 grid, comprising all phonetic sounds articulatable by humanoids, plus mathematical operators and symbols for chemical elements and so on. Symbols can be rotated and mirrored, and similar symbols share phonetic properties. Further complexity can be achieved by increasing the 9-bit grid to a 16, 25, or larger size grid, and by cubing the grid into other dimensions. Banks adds that "while Marain was designed to be as quintessentially clear, concise and unambiguous a language as it is within the wit of human and machine to devise" it also "enables the Culture Minds fully to indulge their seemingly congenital predilection of unnecessary obfuscation, wilful contrariness and the fluent generation of utter and profound confusion in others." In spite of all the graphological specification, the lexicogrammatical workings of Marain remain hidden in the nature of the advanced technological utopia that is the Culture.

Some neographies, though, are worked out far beyond the requirements of the fictional world they occupy. The most famous are J.R.R. Tolkien's numerous constructed languages for the races of Men, Elves, Dwarves, Orcs and others in his fictional world most well-known through *The lord of the rings* series. Tolkien's invented languages include Telerin (the language of the Sea-Elves), Doriathrin (the mother tongue of Lúthien), Nandorin (the Green-Elven language), Khuzdul (the secret language of the Dwarves), Entish (the language of the tree-creatures, the Ents), Orkish, Adûnaic (the vernacular of Númenor), and the two most well-known and most developed, Quenya and Sindarin. Several of Tolkien's languages only have a few words and grammatical features (Entish is merely described as very long-winded and impossible to render in writing), but the last two are strongly reminiscent of Finnish and Welsh, and are fully worked out grammatically with large vocabularies. Quenya is High Elvish and Sindarin is Grey-Elven, with their own histories detailing their dialectal divergence and ultimately different linguistic patterns. Both have elaborate grammatical systems with a complex tense and aspect system, case inflections, and declensions, and rules for syntactic sequencing. Quenya, in particular, was designed for its aesthetic appeal, with Latin sounds blending with the Finnish-derived ones to reduce the number of consonants and so make the language sound more euphonious.

Clearly, Tolkien's invention becomes a scholarly activity beyond the artistic purpose of the novels, though it also adds a richness to the texture of his imagined worlds. The success of these inventions can be judged by the numbers of ardent enthusiasts still engaged in Elvish philology and in trying to piece together the development of Tolkien's linguistic imagination. These activities have also developed far beyond the boundaries of the fictional world in which they originated; and while they are not anywhere near natural languages, the Elvish languages approach a

degree of richness that raise them towards the level of auxiliary languages.

Functions and Assumptions of Invented Languages

In general, there are three broad functions of neographies in literature. The first function is elaborative: the language is present to lend an ornamental richness to the imagined landscape, and works essentially as detailed lyrical description. Klingon, in the *Star trek* universe primarily exists for elaborative purposes. Invented by linguist Mark Okrand for the film, Klingon is the language of an aggressive and war-like race and so it is marked by guttural, uvular phonemes that are felt to convey its harshness. Of course, this phonoaesthetic harshness is perceived only relative to English-speaking norms. Klingon has a dictionary and hundreds of speakers, and some Shakespeare plays have been translated. More eccentrically, portions of the Bible have been rendered into Klingon – directly from the Greek and Hebrew!

Secondly, invented language has an indexical function, which is to assist in the creation of the imagined world and increase the reader's sense of its plausibility. This function is indexical because it is one of the main mechanisms by which readers can determine the degree of accessibility between their own world and the imaginary world. A fully-fledged language helps to create a sense of realism and allows readers to build characters through a focalization in their own vernacular. Hergé's *Tintin* stories featured a Balkan country where Syldavian is spoken. This language was not fully developed by Hergé and includes some contradictory features, but it is drawn very clearly from Dutch and German. The language is being used to mark out the sense of difference and foreignness of the country, in relation to the French of Tintin and his readers.

Thirdly, neographies have an emblematic function in the sense that the language itself represents a thematically important idea. Newspeak, Utopian, Marain, and Brian Aldiss's psychedelic language are all thematically important carriers of the political polemic, comment, or satire of the works in which they feature.

Since most writers of literary fiction are not linguists, several of the assumptions on which their neographies are based do not correspond particularly well with modern linguistic orthodoxy. Most noticeable is the Sapir-Whorf hypothesis that seems to be held in its strongest possible form by many writers. Samuel Delany's *Babel-17* (1968) goes beyond Newspeak in imagining a language which, if learned, subverts the speaker's mind and makes them act traitorously without even realizing it. Neal Stephenson's (1993) *Snow crash* similarly imagines an ancient Sumerian language that is close to the linguistic infrastructure of human brains, which can be manipulated by 'neurolinguistic hackers' in order to 'crash' the mind of anyone who reads it. The literalness of the notion of deep structure and the linguistic determinism displayed in these texts is striking.

The norms for evaluations assumed by many English writers are always English language norms, so alienness is indexed by phonetic and lexicogrammatical systems that are exotic only in relation to English. This cultural specificity is not surprising but gives rise to many of the assumptions about phonoaesthetics that are manifest in invented languages.

Finally, it must be remembered that the literary context is an artistic one, where poetic license ought to be allowed. Fictional extrapolations in the dimension of linguistics seem to be evaluated in a more rigid way than comparable speculations in other sciences. It is striking that debates over linguistic patterns adopted by literary writers become very vehement, as if the language itself, even though imaginary in origin, has a status and reality of its own, and a life beyond the world in which it belongs.

See also: Back-Formation; Compound; Creativity in Language; Language Change, Intentional; Language Planning and Policy: Models; Languages for Specific Purposes; Neologisms; Newspeak; Phonetic Transcription: History; Possible Worlds: Stylistic Applications; Sapir, Edward (1884–1939); Semantic Change: Bleaching, Strengthening, Narrowing, Extension; Whorf, Benjamin Lee (1897–1941).

Bibliography

Aldiss B (1969). *Barefoot in the head.* London: Faber and Faber.
Allan J (ed.) (1978). *An introduction to Elvish.* Taunton: Bran's Head Books.
Banks I M (1988). *Consider Phlebas.* London: Orbit.
Banks I M (1991). *The state of the art.* London: Orbit.
Barnes M J E (1974). *Linguistics and languages in science fiction-fantasy.* New York: Arno.
Bear G (1985). *Eon.* London: Grafton.
Burgess A (1962). *A clockwork orange.* London: Heinemann.
Delany S R (1968). *Babel-17.* New York: Ace Books.
Delany S R (1977). *The jewel-hinged jaw: notes on the language of science fiction.* Elizabethtown: Dragon Press.
Delany S R (1984). *Starboard wine: more notes on the language of science fiction.* Pleasantville: Dragon Press.
Gibson W (1984). *Neuromancer.* London: Victor Gollancz.
Hoban R (1980). *Riddley Walker.* London: Jonathan Cape.
LeGuin U K (1982). *The language of the night: essays on fantasy and science fiction.* New York: Berkeley.

Meyers W E (1980). *Aliens and linguists: language study and science fiction.* Athens: University of Georgia Press.
More T (1910). *Utopia.* London: Dent.
Noon J (1993). *Vurt.* London: Pan.
Okrand M (1997). *Klingon for the galactic traveler.* New York: Simon & Schuster.
Orwell G (1948). *1984.* London: Secker and Warburg.
Stephenson N (1993). *Snow crash.* London: Penguin.
Stockwell P (2000). *The poetics of science fiction.* London: Longman.
Whorf B L (1957). *Language, thought, and reality: selected writings of Benjamin Lee Whorf.* Carroll J B (ed.). Cambridge: MIT Technology Press.
Yaguello M (1991). *Lunatic lovers of language: imaginary languages and their creators.* Slater C (trans.). London: Athlone Press.

Iordan, Iorgu (1888–1986)

M C Augusto, University of Utrecht, Utrecht, The Netherlands

Iorgu Iordan was born the son of a modest family, on September 29, 1888 in Tecuci, a Moldavian town in Romania, and died on September 19, 1986. He studied German Language and Literature, and Law, and finished a Ph.D. on Modern Philology at the University of Iasi (1919) with a dissertation on Romanian phonetics.

From his contact in the early 1920s with other linguists (W. Meyer-Lübke, L. Spitzer, Ferdinand Brunot, Antoine Meillet, and Mario Roques, among others) in Berlin, Bonn, and Paris, but especially from the courses he followed with W. Meyer-Lübke, he became acquainted with the comparative method. This approach of the linguistic object he faithfully continued and developed in his career as a teacher and scholar. At the same time, his attention in his classes and research was increasingly being devoted to the "individual aspect of the evolution of the language," following an idealistic methodology à la K. Vossler. His friendship in Bonn with Leo Spitzer, who defended an idealistic stylistics, must have been of great importance to this development.

Once in Paris, the contacts with J. Gilliéron, whose seminars he attended, had a formative impact on Iorgu Iordan and increased his interest in the investigation about the individual and geographic facets of the language.

All these aspects and influences were absorbed into his scientific production. At least four books reflect the linguistic conceptions and guidelines of the prominent linguist that Iorgu Iordan was: *Rumänische Toponamastik* (Bonn-Leipzig, 1924), *Limba romînă actuala. o gramatica a 'greselilor'* (1943, rev. 1948), *Stilistica limbii române* (1944, rev. 1975), and *Introducere în studiul limbilor romanice. Evoluţţia şi starea actuală a lingvisticii romanice* (1932, rev. 1962). The first has contributed in a special way to a better international knowledge of the Romanian language and has promoted its study at the level of the comparative studies of the Romance languages, earning him the epithet of patriarch of Romanian linguistics. The fourth, translated into many languages, has launched Iorgu Iordan as being among the most remarkable linguists. In this book, the author presents a complete, accurate, and objective critical review of the principal currents and its scholars that represent the linguistics scenario in the first half of the 20th century. Regarding the Romance languages, another important study was the reelaboration of the book *Introducere în linguistica romanică* (1957) in collaboration with Maria Manoliu-Manea, which was published with the same title in 1965. The Spanish edition of this book, *Manual de Lingüistica románica*, translated by Manuel Alvar (1972), has been an indispensable item in every bibliography concerning the Romance languages.

His numerous studies on the comparative aspects of the Romance languages did not prevent him from dedicating himself also to the specialized study of Spanish. In this respect, he also was responsible for the creation of the first chair of Spanish Studies in Bucharest.

In addition to his scholarly activities, Iorgu Iordan held various notable administrative positions. These included the presidency of national and international linguistic associations, ambassador (in Russia, 1945–1947), visiting professor, dean of the Faculty of Philology (1947–1950 and 1956–1957), rector of the University of Bucharest, and many others. However, the founding of the Institute of Linguistics of Bucharest, which he ran as director from 1949 to 1952 and later from 1958 to 1970, and the creation of a department concerned with the Romance Languages reveal his capacity and intuition as a scientific administrator. As a director of this institute, Iorgu Iordan was responsible for two major scientific projects: the coordination of the Dictionary of

Meyers W E (1980). *Aliens and linguists: language study and science fiction.* Athens: University of Georgia Press.
More T (1910). *Utopia.* London: Dent.
Noon J (1993). *Vurt.* London: Pan.
Okrand M (1997). *Klingon for the galactic traveler.* New York: Simon & Schuster.
Orwell G (1948). *1984.* London: Secker and Warburg.
Stephenson N (1993). *Snow crash.* London: Penguin.
Stockwell P (2000). *The poetics of science fiction.* London: Longman.
Whorf B L (1957). *Language, thought, and reality: selected writings of Benjamin Lee Whorf.* Carroll J B (ed.). Cambridge: MIT Technology Press.
Yaguello M (1991). *Lunatic lovers of language: imaginary languages and their creators.* Slater C (trans.). London: Athlone Press.

Iordan, Iorgu (1888–1986)

M C Augusto, University of Utrecht, Utrecht, The Netherlands

Iorgu Iordan was born the son of a modest family, on September 29, 1888 in Tecuci, a Moldavian town in Romania, and died on September 19, 1986. He studied German Language and Literature, and Law, and finished a Ph.D. on Modern Philology at the University of Iasi (1919) with a dissertation on Romanian phonetics.

From his contact in the early 1920s with other linguists (W. Meyer-Lübke, L. Spitzer, Ferdinand Brunot, Antoine Meillet, and Mario Roques, among others) in Berlin, Bonn, and Paris, but especially from the courses he followed with W. Meyer-Lübke, he became acquainted with the comparative method. This approach of the linguistic object he faithfully continued and developed in his career as a teacher and scholar. At the same time, his attention in his classes and research was increasingly being devoted to the "individual aspect of the evolution of the language," following an idealistic methodology à la K. Vossler. His friendship in Bonn with Leo Spitzer, who defended an idealistic stylistics, must have been of great importance to this development.

Once in Paris, the contacts with J. Gilliéron, whose seminars he attended, had a formative impact on Iorgu Iordan and increased his interest in the investigation about the individual and geographic facets of the language.

All these aspects and influences were absorbed into his scientific production. At least four books reflect the linguistic conceptions and guidelines of the prominent linguist that Iorgu Iordan was: *Rumänische Toponamastik* (Bonn-Leipzig, 1924), *Limba romînă actuala. o gramatica a* '*greselilor*' (1943, rev. 1948), *Stilistica limbii române* (1944, rev. 1975), and *Introducere în studiul limbilor romanice. Evolutţia şi starea actuală a lingvisticii romanice* (1932, rev. 1962). The first has contributed in a special way to a better international knowledge of the Romanian language and has promoted its study at the level of the comparative studies of the Romance languages, earning him the epithet of patriarch of Romanian linguistics. The fourth, translated into many languages, has launched Iorgu Iordan as being among the most remarkable linguists. In this book, the author presents a complete, accurate, and objective critical review of the principal currents and its scholars that represent the linguistics scenario in the first half of the 20th century. Regarding the Romance languages, another important study was the reelaboration of the book *Introducere în linguistica romanică* (1957) in collaboration with Maria Manoliu-Manea, which was published with the same title in 1965. The Spanish edition of this book, *Manual de Lingüistica románica*, translated by Manuel Alvar (1972), has been an indispensable item in every bibliography concerning the Romance languages.

His numerous studies on the comparative aspects of the Romance languages did not prevent him from dedicating himself also to the specialized study of Spanish. In this respect, he also was responsible for the creation of the first chair of Spanish Studies in Bucharest.

In addition to his scholarly activities, Iorgu Iordan held various notable administrative positions. These included the presidency of national and international linguistic associations, ambassador (in Russia, 1945–1947), visiting professor, dean of the Faculty of Philology (1947–1950 and 1956–1957), rector of the University of Bucharest, and many others. However, the founding of the Institute of Linguistics of Bucharest, which he ran as director from 1949 to 1952 and later from 1958 to 1970, and the creation of a department concerned with the Romance Languages reveal his capacity and intuition as a scientific administrator. As a director of this institute, Iorgu Iordan was responsible for two major scientific projects: the coordination of the Dictionary of

Rumanian, the so-called *Dicţionarul Limbii române al Academiei*, and the elaboration of the *Crestomaţia Romanică* (1962–1972). Many well-known Romanian linguists such as Maria Iliescu, Maria Manoliu-Manea, Sanda Reinheimer, and Marius Sala have collaborated on this last project.

As an international recognition of his scientific value, Iorgu Iordan was made honorary president, from 1968 onward, of the *Societé de Linguistique Romane*, which he cofounded; he was a member of the *Junta Directiva de la Asociación International de Hispanistas*, a collaborator on many Editorial Boards of Linguistic Journals, and *Doctor Honoris Causa* of several universities (Berlin, Gand, Rome, Montpellier, etc.).

As a first-class teacher and researcher, Iorgu Iordan also stood out as a source of knowledge and inspiration for his students. Many of them would become distinguished scholars working all over the world, nevertheless united by a tangible bond between them, once forged by their old teacher.

See also: Brunot, Ferdinand (1860–1938); Contact-Induced Convergence: Typology and Areality; Gilliéron, Jules (1854–1926); Lexicography: Overview; Meillit, Antoine (Paul Jules) (1866–1936); Meyer-Lübke, Wilhelm (1861–1936); Romance Languages; Spanish; Spitzer, Leo (1887–1960); Vossler, Karl (1872–1949).

Bibliography

Cazacu B, Coteanu I, Fischer I, Graur A, Iliescu M, Istrate G, Macrea D, Niculescu A, Paşca St, Petrovici E, Rosetti A, Sala M & Vianu T (1958). *Omagiu lui Iorgu Iordan cu prilejul împlinirii a 70 de ani.* Bucureşti: Editura Academiei.

Iliescu M (1987). 'Iorgu Iordan (1888–1986).' *Zeitschrift für Romanische Philologie 103,* 695–697.

Iordon I (1939). 'Sufixe românesti de origine recentă.' *Buletinul Institutului de Filologie Română Alexandru Philippide 6,* 1–59.

Iordan I (1949). *Influenţe ruseşti asupra limbii române.* Bucureşti.

Iordan I (1950). *Grammatika rumynskogo jazyka.* Perevod so vtorogo rumynskogo izdanija i predislovie prof. S. B. Bernştejna., Moscova: Izdatel'stvo inostrannoj literatury.

Iordan I (1952). *Nume de locuri româneşti în Republica Populară Română.* Bucureşti: Editura Academiei.

Iordan I (1954). *Limba română contemporană.* Bucureşti: Editura Ministerului Învăţământului.

Iordan I (1955). *Limba "eroilor" lui I. L. Caragiale.* Bucureşti: Societatea de Ştiinţe Istorice şi Filodogice.

Iordan I (coord.) (1962–1974). *Crestomaţie romanică* (Vol. I–III). Bucureşti: Editura Academiei.

Iordan I (1962). *Einführung in die Geschichte und Methoden der romanischen Sprachwissenschaft.* Ins Deutsche übertragen, ergänzt und teilweise neu bearbeitet von Werner Bahner. Berlin: Akademie Verlag.

Iordan I (1962). *Lingvistica romanică. Evoluţie. Curente. Metode.* Bucureşti: Editura Academiei.

Iordan I (1963). *Istoria limbii spaniole.* Bucureşti: Editura Didactică şi Pedagogică.

Iordan I (1963). *Toponimia românească.* Bucureşti: Editura Academiei.

Iordan I (1964). *Lexicul limbii române. Le lexique de la langue roumaine.* Bucureşti: Universitatea din Bucureşti Seria *Limbă: 1.*

Iordan I (1964). 'Unele aspecte ale formării cuvintelor în limba română actuală.' *Studü si Cercetări Lingvistice 15(4),* 401–422.

Iordan I & Manoliu M (1965). *Introducere în lingvistica romanică.* Bueureşti: Editura Didactică şi Pedagogică.

Iordan I (1966). *Tendinţe sintactice în limba română actuală.* Bucureşti: Universitatea Bucureşti.

Iordan I (1967). *Lingüística románica. Evolución. Corrientes. Métodos.* Reelaboración parcial y notas de Manuel Alvar. Madrid: Ediciones Alcalá.

Iordan I, Guţu-Romalo V & Niculescu A (1967). *Structura morfologică a limbii române contemporane.* Bucureşti: Editura Ştiinţifică.

Iordan I (1969). 'Note de lexicologie românească.' *Studii si Cercetări Lingvistice 20(1),* 3–11.

Iordan I (1969). *Dos estudios de lingüística románica.* Montevideo: Universidad de la República, Facultad de Humanidades y Ciencias, Departamento de Lingüística.

Iordan I & Orr J (1970). *An Introduction to Romance Linguistics. Its Schools and Scholars.* Revised, with a Supplement Thirty Years on by R Posner. Oxford: Basil Blackwell.

Iordan I (1973). *Introdução à linguistíca românica.* Ferreira J D (trans.). Lisboa: Fundação Calouste Gulbenkian.

Iordan I (1973). *Introduzione alla linguistica romanza.* Cedrini L B (trans.). Torino: Einaudi.

Sala M (1967). 'Iorgu Iordan.' *Boletín de Filologí Española 22–23,* 3–15.

Sala M (1989). 'Iorgu Iordan (1888–1986).' *Revista de Filología Española 69,* 349–352.

used for *θ* and *t*, respectively, but in a few manuscripts ⟨ṭ⟩ is *t* and ⟨ṭ⟩ is *θ*. Vowels are marked using the Syriac pointing system, which makes this the only variety of Sogdian script in which the crucial distinctions between *a-* and *ā-*, *ī̆* and *ē̆*, *ū̆* and *ō̆* are indicated in the script. Short *a* is noted also with other consonants, e.g., *wa*.

See also: Aramaic and Syriac; Avestan; Iranian Languages; Persian, Modern; Persian, Old.

Bibliography

Hoffmann K & Forssman B (1996). *Avestische Laut-und Flexionslehre*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Hoffmann K & Narten J (1989). *Der Sasanidische Archetypus*. Untersuchungen zu Schreibung und Lautgestalt des Avestischen. Wiesbaden: Reichert..

Kent R G (1953). *Old Persian grammar, texts, lexicon* (2nd rev. edn.). New Haven, CT: American Oriental Society.

Naveh J & Shaked S (eds.) (2005). *Ancient Aramaic documents from Bactria (fourth century, B.C.E.)*. Corpus Inscriptionum Iranicarum (Part I, Vol. 5, Texts 2). London: Corpus Inscriptionum Iranicarum.

Sims-Williams N (1985). *The Christian Sogdian manuscript C2*. Berliner Turfantexte 12. Berlin: Akademie-Verlag.

Skjærvø P O (1995). 'Aramaic in Iran.' *ARAM 7* (*Palmyra and the Aramaeans*) 283–318.

Skjærvø P O (1996). 'Aramaic scripts for Iranian languages.' In Daniels P T & Bright W (eds.) *The world's writing systems*. New York: Oxford University Press. 515–535.

Stronach D (1990). 'On the genesis of the Old Persian cuneiform script.' In Vallat F (ed.) *Contribution à l'histoire de l'Iran. Mélanges offerts à Jean Perrot*. Paris: Recherche sur les civilisations. 195–203.

Weber D (2003). *Berliner Papyri, Pergamente und Leinenfragmente in mittelpersischer Sprache* unter Mitarbeit von W. Brashear†. Ägyptische Urkunden aus den Staatlichen Museen Berlin. Pahlavi-Dokumente. Corpus Inscriptionum Iranicarum (Part III, vols 4–5. London: School of Oriental and African Studies. 168–194.

Iran: Language Situation

B Ingham, School of Oriental and African Studies, University of London, UK

Main Language Families

Three main languages families are spoken in Iran, covering discrete areas, but overlapping in places. This is facilitated by the thinness of the population in parts. The majority group in terms of number of speakers and geographical area is the Iranian group of the Indo-European family, spoken as original local languages over most of the country except in Azerbaijan and Khuzistan. The national language Persian (Farsi) was originally a dialect of the region of Fars and has a written history in the present alphabet of about 1000 years. It is spoken in all the main cities and in the northeast. Kindred dialects are those of Fars and of the Lur tribes of the south. Other Iranian varieties include (Schmitt, 1989: 294–295) (i) Kurdish in the west, (ii) Tati and Talishi in the north west, (iii) Gilaki and Mazandarani north of the Elburz mountains, (iv) Central dialects around Hamadan, Tehran, Isfahan, and Yazd, and (v) Baluchi in the southeast. Most of these are quite distinct from Farsi and not mutually intelligible with it. However, only Kurdish, Baluchi, and to a lesser extent Luri are usually accorded separate language status, because of the political power of their speakers. Kurdish and Baluchi are written languages. The next in size is the Turkic group, represented by Azeri in the north west (Azerbaijan), by Uzbek, spoken by the Turkomans in the northeast, by large numbers of Turkish speakers in Fars in the southwest, of particular importance being the tribal Qashqai and Khamse and in other localities, spreading round from northeast to southwest interspersed with the Iranian-speaking population. The two groups overlap considerably, this sometimes resulting from resettlement by previous governments. The third, the Semitic group, is represented by Arabic spoken in Khuzistan in the southwest and at various localities along the Gulf Coast, the Makran Coast and possibly also by some of the nomadic population of Fars among the Khamse. There are also Arabic speaking populations in the northeast around Sarakhs and in the east around Birjand. The above are all non-minority languages, spoken over large areas by native populations spanning the borders of Iran and neighboring states. Persian spreads into Afghanistan (known there as Darri) and Tadjikistan (known there as Tadjik). Turkic spreads into Turkey and Azerbaijan in the west, into Turkmenistan and other central Asian states in the northeast, and into Afghanistan in the east. Arabic spreads westward into Iraq. The population of the Persian Gulf islands speak Arabic or Persian or are bilingual.

Minority Languages

Minority languages, spoken by mainly, but not exclusively urban ethnic minorities include Armenian, an Indo-European language spoken by the Christian Armenians, Georgian, a Caucasian language spoken by the Georgian Christians, Syriac, a Semitic language spoken by the Nestorian Christian Assyrians (āšūrīhā) an Arabic dialect spoken by some of the Jewish population, akin to the *qiltu* dialects of Northern Iraq (Blanc, 1964: 7), and an Indo-European variety spoken by some of the Ghorbati or Kawli, the Gypsies of Iran, which may be a dialect of Persian.

Status of the Different Languages

Persian is the main medium of instruction in schools and is the language of the media and the administration. Turkic, Kurdish, Arabic, and Baluchi, however, are of considerable local importance, being the language of rural elites. The case of Arabic is slightly different since Khuzistan, being the major oil region, has undergone considerable immigration and urbanization, so that Arabic is heard more often in rural areas. Here we intend the spoken dialect, which is akin to the southern *gilit* type of Southern Iraq (Blanc, 1964: 7). The classical, literary language is accorded high esteem especially by the present regime because of its association with Islam. Persian is understood universally in the towns and widely in rural areas, but even after the efforts of the Pahlavi regime to universalize it by means of education and the media, it is not everywhere spoken or even understood in rural areas. Note also that the appellation Arab or Turk does not necessarily imply a language affinity, but may be only a political one, as with some of the 'Arab' tribes of Fars and Afghanistan who speak Persian or Turkish (Barth, 1961: 131). The main languages taught in schools are Classical Arabic, taught as a language of liturgy and scholarship, but not as a spoken tongue, and English. At an earlier stage, French was more common.

See also: Indo–Iranian; Persian, Modern; Turkmen.

Bibliography

Barth F (1961). *Nomads of South Persia.* Oslo: Oslo University Press.

Blanc H (1964). *Communal dialects in Baghdad.* Cambridge, Mass: Harvard University Press.

Christensen A (1930–1935). *Contributions a la dialectologie iranienne I and II, Det Kgl. Danske Videnskabernes Selskab.* Copenhagen: Bianco Lunos Bogtrykkeri A/S and Levin and Munksgaard. 16 ii, 21 iii.

Coon S (1952). *Caravan, the story of the Middle East.* London: Jonathan Cape.

Ingham B (1973). 'Urban and rural Arabic in Khuzistan.' *Bulletin of the School of Oriental and African Studies 36,* iii, 533–553.

Ingham B (1976). 'Regional and social factors in the dialect geography of Southern Iraq and Khuzistan.' *Bulletin of the School of Oriental and African Studies 39,* i, 62–82.

Ingham B (1982). *North east Arabian dialects.* London: Kegan Paul International.

Lambton A K S (1938). *Three Persian dialects.* London: Royal Asiatic Society.

Lorimer D L R (1921). *The Phonology of the Bakhtiari, Badakhsani and Madaglashti dialects of Modern Persian.* London: Royal Asiatic Society.

Mann O (1904). *Kurze Skizze der Lurdialekte, Sitzungsberichte der Akademie der Wissenschaften zu Berlin 39.* Berlin: Druck und Verlag von Georg Reimer.

Mann O (1909). *Die Tajik Mundarten der Provinz Fars, Kurdisch-Persische Forschungen I.* Berlin: Druck und Verlag von Georg Reimer.

Mann O (1910). *Die Mundarten Lur Stamme im sudwestlichen Persien, Kurdisch-Persische Forschungen II.* Berlin: Druck und Verlag von Georg Reimer.

Menges K H (1951). 'Research in the Turkic dialect of Iran.' *Oriens 4,* 273–279.

Oberling P (1974). *The Qashqai nomads of Fars.* The Hague: Mouton.

Oranskij I M (1975). *Die neuiranische Sprachen der Sowjetunion.* The Hague: Mouton.

Schmitt R (ed.) (1989). *Compendium linguarum Iranicarum.* Wiesbaden: Verlag.

Seeger U (2002). 'Zwei Texte im Dialekt der Araber von Chorasan.' In Arnold W & Bobzin H (eds.) *Sprich doch mit deinen Knechten aramaisch, wir verstehen es! 60 Beitrage zur Semitistik. Festschrift fur Otto Jastrow zum 60. Geburtstag.* Wiesbaden: Otto Harrassowitz. 299–309.

Sykes P (1902). 'Anthropological notes on southern Persia.' *Journal of the Anthropological Institute 32,* 339–352.

Sykes P (1906). 'The Gypsies of Persia.' *Journal of the Anthropological Institute 36,* 302–311.

Windfuhr G (forthcoming). *The Iranian languages.* London: Routledge.

Iran: Scripts, Old Persian, Aramaic, Avestan

P O Skjærvø, Harvard University, Cambridge, MA, USA

Writing systems for Iranian languages include cuneiform (Old Persian); scripts descended from 'imperial' Aramaic (earliest monuments of Chorasmian, Sogdian, Parthian, and Middle Persian, as well as Avestan, derived from the Middle Persian script); two Syriac scripts, Estrangelo or 'Manichean' (Middle Persian, Parthian, Bactrian, Sogdian, Modern Persian) and 'Nestorian' (Christian Sogdian, Modern Persian); Hebrew (Modern Persian and local Persian dialects: Judeo-Persian); Arabic (Modern Persian, Chorasmian, Kurdish, Pashto, Balochi); Brahmi (Khotanese, Tumshuqese, Sogdian); Greek (Bactrian); Cyrillic (Ossetic, Tajik); Georgian (Ossetic); and Latin (Kurdish, Ossetic).

Old Persian Cuneiform

Old Persian is written in a cuneiform script, but with only a small number of signs. There are 3 vowel signs ⟨a, i, u⟩, 33 signs of the form 'consonant + vowel' (CV), 8 logograms, a word separator, and numerals. The CV signs contain the vowels *a*, *i*, or *u* ⟨Ca, Ci, Cu⟩, but the set of ⟨Ci⟩ and ⟨Cu⟩ signs is not complete. There are only 4 ⟨Ci⟩ signs and 7 ⟨Cu⟩ signs, and most of the ⟨Ca⟩ signs also have to serve for the missing ⟨Ci⟩ and ⟨Cu⟩ signs. Thus, crucial morphological distinctions such as the 3rd singular endings active *-tiy* and middle *-taiy*, both ⟨-ta-i-ya⟩, are not indicated. The ⟨Ca⟩ signs are also used where a consonant is followed by no vowel, before other consonants, or in word-final position (only *m*, *r*, and š).

The vowel signs ⟨a, i, u⟩ are used initially (when the word begins with a vowel). Medially, ⟨a⟩ is used to write long *ā*. The signs ⟨i⟩ and ⟨u⟩ are used medially to write short *i* and *u* and long *ī* and *ū*. They are preceded by the ⟨Ci⟩ and ⟨Cu⟩ signs whenever these exist. Very rarely in the attested corpus, ⟨Ci⟩ and ⟨Cu⟩ are used alone to indicate consonant plus *i* or *u*. In final position, they are followed by ⟨ya⟩ and ⟨va⟩ (*-ĭ̄y*, *-ŭ̄v*).

Certain sounds that are expected from the etymology of the words are not written in the Old Persian script, notably *h* before *u* and *m* and nasals before consonants, although they may be present both in transcriptions into other languages and in later stages of Persian:

h: *Ahuramazdā* is spelled ⟨a-u-ra-ma-za-da-a⟩, but ⟨a-ḫu-ru-ma-az-da-'⟩ in Akkadian; the Middle Persian form is *Ōhrmazd*; *ahmiy* 'I am' is spelled both ⟨a-mi-i-ya⟩ and ⟨a-ha-mi-i-ya⟩; with *tauhmā* ⟨ta-u-ma-a⟩ 'family', cf. Av. *taoxman*-'offspring', NPers. *toxm* 'seed'; and with *ahrīka* ⟨a-ra-i-ka⟩ 'on the side of the Evil one', cf. Mid. Pers. *ahrīmen* 'the Evil one';

nasals: *Kambŭ̄jiya-* is spelled ⟨ka-ba-u-ji-i-ya⟩, but Elamite ⟨kán-bu-zi-ia⟩, Akkadian ⟨kam-bu-zi-ia⟩, Aramaic ⟨kmbwzy⟩, Greek *Kambúsēs*; with *bandaka* 'bondsman', spelled ⟨ba-da-ka⟩, cf. Middle Persian *bandag* 'servant'.

The script was therefore not quite phonemic. It is also not certain to what extent the written consonants were unambiguous; for instance, diachrony and synchrony suggest that ⟨ja⟩ may be for both /ǰ/ and /ž/, if these were separate phonemes. (See Table 1.)

Numerals

The numerals are simple: the unit is 𐏑 '1', 𐏒𐏑 is '3', etc.; 'ten' is 𐏓 (𐏓𐏒𐏒𐏒𐏒𐏑 '19', etc.); and '20' is' 𐏔 (𐏔𐏒𐏒𐏒𐏑 '27', etc.). The only higher numeral attested is 𐏕 𐏔 '120'.

Scripts Derived from Aramaic

All the Iranian scripts derived from Aramaic express vowels other than short *a* by means of *matres lectionis* as follows: ⟨'⟩ initially = short or long *ă̄-* (long initial *ā-* is often written ⟨''-⟩); between vowels = long *ā* (in the Sogdian script also short *a*).

⟨y⟩ and ⟨w⟩: between vowels = short or long *ĭ̄* or *ĕ̄* and *ŭ̄* and *ŏ̄*, respectively; in the Sogdian script ⟨'y⟩ and ⟨'w⟩ is used for *āy* and *āw* (also ⟨''y, ''w⟩) or instead of ⟨y, w⟩; in the Manichean script initial ⟨ʿy-⟩ *ĭ̄-* may be distinguished from ⟨'y⟩ *ē-*. In Christian Sogdian in the Nestorian Syriac script, vowel quality/quantity is sometimes indicated by vowel marks.

Scripts Derived from Imperial Aramaic

During the Achaemenid period (549–330 B.C.E.), Aramaic ('Imperial Aramaic') was used as the chancellery language, and the Aramaic script was subsequently, at the latest by the Parthian period (ca. 210 B.C.E.–224 C.E.), adopted by several Iranian states: Parthia, the local Persian dynasty, Chorasmia, and Sogdiana. There is a small corpus of Aramaic documents from Bactria from the 4th century B.C.E., and Aramaic was also used in the edicts of King Aśoka (3rd century B.C.E.) written in ancient Bactria and Gandhara (northern Pakistan).

From the early Parthian period (2nd–1st centuries B.C.E.) there are a number of inscriptions, some of which are clearly in Iranian languages, presumably Parthian (thus, e.g., the wine receipts from Nisa, east of the Caspian Sea), although the fact is obscured by

Table 1 The Old Persian syllabary

Ca		Ci		Cu	
𐎠	a	𐎡	i	𐎢	u
𐎲	ba				
𐎨	ca				
𐏂	ça				
𐎭	da	𐎮	di	𐎯	du
𐎳	fa				
𐎥	ga			𐎦	gu
𐏃	ha				
𐎩	ja	𐎪	ji		
𐎣	ka			𐎤	ku
𐎾	la				
𐎶	ma	𐎷	mi	𐎸	mu
𐎴	na			𐎵	nu
𐎱	pa				
𐎼	ra			𐎽	ru
𐎿	sa				
𐏁	ša				
𐎫	ta			𐎬	tu
𐎰	θa				
𐎺	va	𐎻	vi		
𐎧	xa				
𐎹	ya				
𐏀	za				
𐏐 𐏐 :	(word dividers)				

Logograms

𐏋	XŠ = xšāyaθiya	𐏏	BU = būmiš
𐏌	DH1 = dahạyāuš	𐏈	AM1 = Auramazdā
𐏍	DH2 = dahạyāuš	𐏉	AM2 = Auramazdā
𐏎	BG = baga	𐏊	AMha = Auramazdāha

the extensive use of aramaeograms (heterograms, ideograms). Others are more likely to be in faulty bureaucratic Aramaic. By the 2nd century C.E., the official Parthian script had reached its standard form.

The Middle Persian variant, seen on the coins of pre-Sasanian kings, evolved in southern Iran in the late Parthian period and became the official script of the Sasanian empire (224–651 C.E.), seen in royal and private inscriptions from the 3rd to the 5th centuries. Variants of this script are found on the coins of the Sasanian kings. A more cursive variant of the Middle Persian script used for writing on paper is the so-called Psalter script, known from a fragmentary manuscript of the Psalms of David found in Chinese Turkestan. This script developed via a simplified epigraphic variant (known from funerary inscriptions) into the so-called (Book) Pahlavi script, used in the Zoroastrian books. The final stage is the almost indecipherable script used in the Pahlavi papyri and parchments of the late Sasanian early Islamic period.

The Avestan script is a composite script made on the basis of the Psalter and Book Pahlavi scripts.

The Chorasmian script is known from a few inscriptions; it was later replaced by the Arabic script.

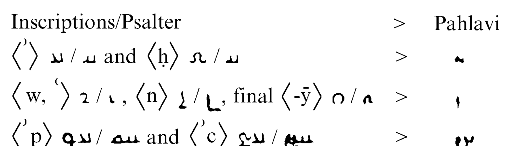
The Sogdian script is known in three principal forms. The earliest form is used in letters written on paper and dating from the early 4th century C.E. and in a few short inscriptions from northern Pakistan. The common Sogdian script is a cursive variant found in secular documents and in Buddhist and Manichean manuscripts, of which two stages are distinguished: the 'formal' or 'sutra script,' and the more ambiguous 'Uighur' variant. (See Table 2.)

Use of Aramaic Letters

The Aramaic alphabet contained letters for numerous sounds not found in Iranian. Some of these were adopted to write sounds that were not found in Aramaic (⟨ṣ⟩ for Iranian *č* [tʃ] and ⟨ḥ⟩ for the fricative velar /x/; ⟨l⟩ in Sogdian for /δ/), but even so the Aramaic alphabet contained more letters than were needed, and ⟨ṭ, ʿ, q⟩ were never used in writing Iranian words (except with numerical values; differently in the Syriac scripts). In Middle Persian, ⟨l⟩ is used more often than ⟨r⟩ (= ⟨w⟩) to write *r*. In Aramaic documents from the mid-Parthian era, ⟨ḥ⟩ and ⟨ẖ⟩ are being confused, and, in Parthian and Middle Persian, ⟨ẖ⟩ is confined to final position in aramaeograms while ⟨ḥ⟩ is written elsewhere, regardless of their etymologies. Aramaic ⟨ẖ⟩ was also used in Parthian and Sogdian for the feminine ending -*ă*, which resembled the Aramaic ending -*ā*, but this became a mere convention when the ending was lost both in Parthian and in Sogdian in 'heavy' stems.

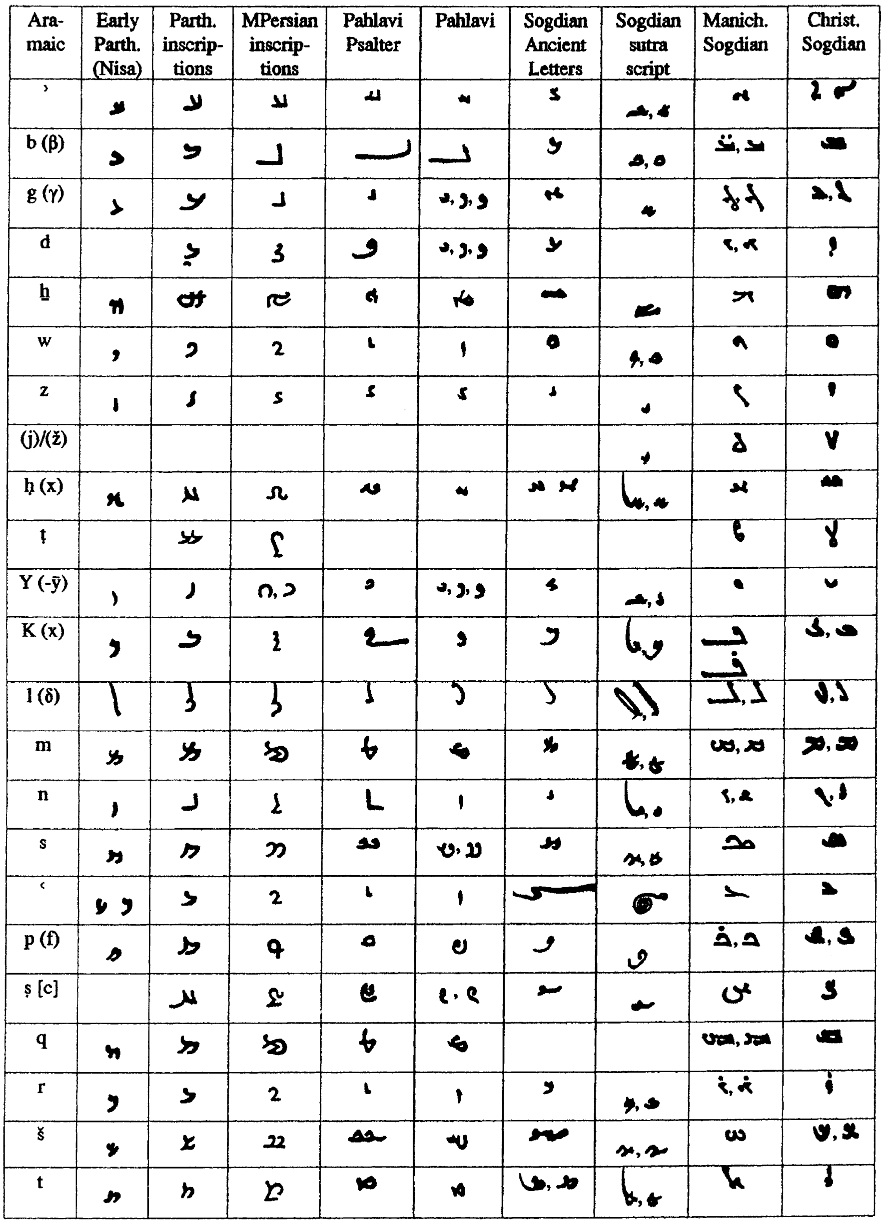
The Middle Persian Script

Typical of the development of the Middle Persian script is the ever-increasing merger of letters into the same shape, for example:



Out of the 22 distinct letters of the Aramaic alphabet and the 18 of the inscriptions and the *Psalter*, the Pahlavi alphabet distinguishes only 12 letters, which have to do duty for 21. As we see from Table 3, some of these letters have at least two different orthographic values (Psalter forms in parenthesis). As the script

Table 2 The Main Iranian scripts developed from Aramaic and Syriac



The table contains a selection of initial, medial, and final variants.

is strongly cursive, there are numerous positional variants.

Of these, ⟨ʿ⟩ and ⟨q⟩, as well as ⟨h̲⟩ (= ⟨m + w, n⟩) are found only in aramaeograms, transliterated as ⟨O, Q, E⟩, and ⟨Q⟩ is found only in ⟨QDM⟩ *abar* 'on'.

In Indian manuscripts, ⟨š⟩ is replaced by ⟨y'⟩.

To remedy the ambiguities, diacritical marks are sometimes used (not always etymologically correct): ⟨d⟩, ⟨g⟩, ⟨y⟩, ⟨j̆⟩ (initial).

In practice, therefore, most of the time Pahlavi has to be read word by word or morpheme by morpheme, rather than letter by letter. For instance, *Ohrmazd*, inscr. ⟨'whrmzd⟩ became ⟨'whrmzd⟩ = ⟨'w'wm'⟩ = ⟨'nhwm'⟩ (hence the traditional Parsi pronunciation *anhuma*). The word *yazdān* 'god(s)', inscr. became ⟨yzd̲'n'⟩ = ⟨y''ww⟩, which was distinguished from *dēwān* 'demons', inscr. , Pahl. ⟨ŠDYAn'⟩ = ⟨š'''ww⟩ (= ⟨y''''ww⟩ with ⟨y'⟩ for ⟨š⟩) by the number of loops,

Table 3 The 12 distinct Pahlavi letters

	ʾ ( ), h ( )		z		p
	b		k		c
or	d ( ), g ( ), y ( )		l		š
	w, ʿ, r ( ), n ( )		m, q		t

which in the manuscripts are sometimes multiplied even further (, etc.). The word *mǎzdēsn* 'Mazdayasnian' is usually written like *mahist* ⟨mhyst'⟩ 'greatest' with ⟨yt⟩ for ⟨sn⟩.

The name of the devil, Ahrimen, is sometimes written up-side-down .

The Avestan Script

The Avestan script was invented probably sometime in the 6th century C.E. in order to write down the Avestan texts, which until then had been transmitted only orally. It is an alphabetical script, in which each allophone of the language as recited at the time corresponds to one letter. The earliest known forms of the Avestan and Pahlavi scripts are those of the 13/14th-century manuscripts and probably represent later developments of the original script. Some letters may have been added in the course of time. (See Table 4.)

The alphabet was based upon the official Pahlavi script, using diacritical marks (e.g., ⟨ī⟩ and ⟨ū⟩ from ⟨i⟩ and ⟨u⟩ and some letter forms from different stages of the Pahlavi script (e.g., Psalter ⟨d⟩). It has also been suggested that ⟨ə⟩ was adopted from Greek ε.

By the time of the earliest manuscripts, the original value of some of the letters was no longer known, notably, the distribution of the three sibilants ⟨š, ş, š́⟩ is not consistent in the manuscripts, and their values have been inferred on linguistic grounds. Several letters are found only in certain manuscripts (e.g., ⟨ń⟩ and ⟨ġ⟩). ⟨y, Y⟩, and ⟨v⟩ are used only initially, while ⟨uu⟩ and ⟨ii⟩ are used medially.

Syriac Scripts

Two forms of Syriac script were used for Iranian languages. The Manichean script, a modified estrangelo (sometimes, but probably apocryphally, ascribed to Mani), was used for Middle Persian, Parthian, Bactrian, Sogdian, and Persian (it was also used for the non-Iranian Tokharian and Old Turkish).

Table 4 The Avestan alphabet

a	ā	ạ̊	å	ą
i	ī	e	ē	ə
u	ū	o	ō	ə̄
b	β	p	f	m
				^{h}m
d	δ	t	θ	n
			ṯ	
g	γ	k	x	ŋ
ġ				
j		c		ń
Y-			x́	ŋ́
y-				
v-			x^v	$ŋ^v$
s	z		h	ṇ
š	ž			
š́				
ṣ̌				

The Nestorian script was used by the Christians at Turfan.

The Manichean Script

As in the Sogdian script, Syriac ⟨l⟩ was used to write *δ*. Diacritics were used to express specifically Iranian sounds, e.g., superscript double dots over ⟨b⟩ = ⟨*β*⟩, ⟨g⟩ with a loop = ⟨*γ*⟩. The only new letter, ⟨ ⟩, is commonly transliterated as ⟨ǰ⟩, while ⟨z⟩ with double dots is transliterated as ⟨ž⟩.

Initial sibilant + stop took a prosthetic vowel written ⟨ʿ⟩ (Mid. Pers. ⟨ʿsp'ẖ⟩ = *əspāh* 'army', Parth. ⟨ʿsp'd⟩ = *əspāδ*). Words with *ē* or *ī* and *ō* or *ū* may be distinguished by the use of ⟨ʿ⟩ or ⟨'⟩ (e.g., Mid. Pers. ⟨'yg⟩ *ēg* 'then' versus ⟨ʿyg⟩ *ī(g)*, relative particle; ⟨zwr⟩ *zōr* 'strength' versus ⟨zʿwr⟩ *zūr* 'deception').

Sometimes, letters marking vowels are left out and the omission indicated by super/subscript dots (e.g., ⟨w⟩ with double dots for ⟨'wd⟩ *ud* 'and' and ⟨š⟩ with double dots for ⟨'wš⟩ *u-š* 'and he/by him/his'). In the Bactrian Manichean script, the double dots were used to distinguish the palatal *ś* and *ź* from the presumably palato-alveolar *š* and ž.

The Nestorian Script

This script included vowel marks and contains a new letter (corresponding to Man. , but commonly transliterated as ⟨ž⟩). The letters ⟨t⟩ and ⟨ṭ⟩ are most often

used for θ and *t*, respectively, but in a few manuscripts ⟨ṭ⟩ is *t* and ⟨ṯ⟩ is θ. Vowels are marked using the Syriac pointing system, which makes this the only variety of Sogdian script in which the crucial distinctions between *a-* and *ā-*, *ĭ* and *ĕ*, *ŭ* and *ŏ* are indicated in the script. Short *a* is noted also with other consonants, e.g., *wa*.

See also: Aramaic and Syriac; Avestan; Iranian Languages; Persian, Modern; Persian, Old.

Bibliography

Hoffmann K & Forssman B (1996). *Avestische Laut-und Flexionslehre*. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Hoffmann K & Narten J (1989). *Der Sasanidische Archetypus*. Untersuchungen zu Schreibung und Lautgestalt des Avestischen. Wiesbaden: Reichert..

Kent R G (1953). *Old Persian grammar, texts, lexicon* (2nd rev. edn.). New Haven, CT: American Oriental Society.

Naveh J & Shaked S (eds.) (2005). *Ancient Aramaic documents from Bactria (fourth century, B.C.E.)*. Corpus Inscriptionum Iranicarum (Part I, Vol. 5, Texts 2). London: Corpus Inscriptionum Iranicarum.

Sims-Williams N (1985). *The Christian Sogdian manuscript C2*. Berliner Turfantexte 12. Berlin: Akademie-Verlag.

Skjærvø P O (1995). 'Aramaic in Iran.' *ARAM 7* (*Palmyra and the Aramaeans*) 283–318.

Skjærvø P O (1996). 'Aramaic scripts for Iranian languages.' In Daniels P T & Bright W (eds.) *The world's writing systems*. New York: Oxford University Press. 515–535.

Stronach D (1990). 'On the genesis of the Old Persian cuneiform script.' In Vallat F (ed.) *Contribution à l'histoire de l'Iran. Mélanges offerts à Jean Perrot*. Paris: Recherche sur les civilisations. 195–203.

Weber D (2003). *Berliner Papyri, Pergamente und Leinenfragmente in mittelpersischer Sprache* unter Mitarbeit von W. Brashear†. Ägyptische Urkunden aus den Staatlichen Museen Berlin. Pahlavi-Dokumente. Corpus Inscriptionum Iranicarum (Part III, vols 4–5. London: School of Oriental and African Studies. 168–194.

Iran: Language Situation

B Ingham, School of Oriental and African Studies, University of London, UK

Main Language Families

Three main languages families are spoken in Iran, covering discrete areas, but overlapping in places. This is facilitated by the thinness of the population in parts. The majority group in terms of number of speakers and geographical area is the Iranian group of the Indo-European family, spoken as original local languages over most of the country except in Azerbaijan and Khuzistan. The national language Persian (Farsi) was originally a dialect of the region of Fars and has a written history in the present alphabet of about 1000 years. It is spoken in all the main cities and in the northeast. Kindred dialects are those of Fars and of the Lur tribes of the south. Other Iranian varieties include (Schmitt, 1989: 294–295) (i) Kurdish in the west, (ii) Tati and Talishi in the north west, (iii) Gilaki and Mazandarani north of the Elburz mountains, (iv) Central dialects around Hamadan, Tehran, Isfahan, and Yazd, and (v) Baluchi in the southeast. Most of these are quite distinct from Farsi and not mutually intelligible with it. However, only Kurdish, Baluchi, and to a lesser extent Luri are usually accorded separate language status, because of the political power of their speakers. Kurdish and Baluchi are written languages. The next in size is the Turkic group, represented by Azeri in the north west (Azerbaijan), by Uzbek, spoken by the Turkomans in the northeast, by large numbers of Turkish speakers in Fars in the southwest, of particular importance being the tribal Qashqai and Khamse and in other localities, spreading round from northeast to southwest interspersed with the Iranian-speaking population. The two groups overlap considerably, this sometimes resulting from resettlement by previous governments. The third, the Semitic group, is represented by Arabic spoken in Khuzistan in the southwest and at various localities along the Gulf Coast, the Makran Coast and possibly also by some of the nomadic population of Fars among the Khamse. There are also Arabic speaking populations in the northeast around Sarakhs and in the east around Birjand. The above are all non-minority languages, spoken over large areas by native populations spanning the borders of Iran and neighboring states. Persian spreads into Afghanistan (known there as Darri) and Tadjikistan (known there as Tadjik). Turkic spreads into Turkey and Azerbaijan in the west, into Turkmenistan and other central Asian states in the northeast, and into Afghanistan in the east. Arabic spreads westward into Iraq. The population of the Persian Gulf islands speak Arabic or Persian or are bilingual.

Iranian Languages

P O Skjærvø, Harvard University, Cambridge, MA

Iranian languages have been spoken for 3000 to 4000 years in various parts of southern Russia and the Caucasus, Central Asian republics, Xinjiang, Afghanistan, Pakistan, Iran, Iraq, and Turkey, as well as in the diaspora. The language spoken by the largest number of people today is (New) Persian (*see* **Persian, Modern**), or Farsi. Iranian languages are closely related to the Indo–Aryan (*see* **Indo–Aryan Languages**) languages, with which they constitute the Indo–Iranian (*see* **Indo–Iranian**) subgroup of the Indo–European language family, to which most European languages also belong.

The Iranian languages are known from three main periods, commonly classified as Old, Middle, and New (modern) Iranian. Historically, this division corresponds roughly to (1) the pre-Achaemenid and Achaemenid period (down to ca. 300 B.C.E.), (2) the period down to the Arabic invasion of Iran and the spread of Islam (7th century C.E. in the west, up to 11th and 12th centuries C.E. in the east), and (3) the Islamic period.

Iranian languages are written in a large variety of ancient and modern scripts (*see* **Iran: Scripts, Old Persian, Aramaic, Avestan**).

Old Iranian Languages

Avestan (*see* **Avestan**), the language of the Avesta, was spoken in areas of Central Asia, Afghanistan, and eastern Iran from the 2nd millennium till about 500 B.C.E. Two stages of the language are known, the older of which is comparable to the oldest Indic seen in the Rigveda and was therefore probably spoken about the mid 2nd millennium B.C.E., while the younger is comparable to Old Persian and was probably spoken ca. 1000–500 B.C.E. The Avestan texts were transmitted orally for 1000 to 2000 years and written down only in about the 6th century C.E.; the earliest manuscripts date from the 13th century. Its phonetic form is therefore that of the latest oral transmitters. The present Avestan corpus is relatively small.

Old Persian (*see* **Persian, Old**), ancestor of Middle and New Persian, was spoken in southwestern Iran in the first half of the 1st millennium, till about 400 B.C.E. It was the official language of the Achaemenid dynasty and was written in a cuneiform script. The Old Persian corpus is quite small.

Numerous words in the Old Persian inscriptions have a phonetic form different from what is considered to be genuinely Old Persian. It is assumed these are from Median, a language spoken to the north of Old Persian, of which no texts survive.

Middle Iranian Languages

Khotanese (*see* **Khotanese**), spoken in the kingdom of Khotan in southwestern Xinjiang, is known from a variety of texts dating from about the 6th century to the end of the 10th century C.E. Tumshuqese, spoken in Kucha in northwestern Xinjiang, is known from the same type of sources, but much less well. These two languages were written in the southern and northern variants of Brahmi, respectively.

Sogdian (*see* **Sogdian**), spoken in Sogdiana, modern Central Asia, is known from texts dating from the 4th to the 10th century. It was written in Sogdian, Manichean, Syriac, and, occasionally, northern Brahmi scripts.

Chorasmian (*see* **Chorasmian**), spoken in the Chorasmian state along the upper Oxus/Syr Darja as late as the 14th century, is known from inscriptions, coins, and interlinear glosses in two Arabic works, a legal work and a dictionary. It was originally written in Chorasmian, later in Arabic script.

Bactrian (*see* **Bactrian**), spoken in the Greco–Bactrian kingdom founded by soldiers of Alexander the Great in northern Afghanistan, is best known from royal inscriptions and private documents dating from the 2nd to the 8th century. It was written in lapidary and cursive Greek and in Manichean scripts.

Parthian, spoken in the Parthian kingdom to the northeast of the Caspian Sea, is known from inscriptions, letters, economic and legal documents, and Manichean texts dating from about the 1st century B.C.E. to the 10th century. It was written in Parthian and Manichean scripts.

Middle Persian (also Pahlavi; *see* **Pahlavi**), was spoken in southwestern Iran and became the official language of the Sasanian dynasty. It is known from a large variety of texts, notably Zoroastrian texts, dating from about the 1st century B.C.E. to the 13th century, although it had been replaced by modern Persian as a spoken language by about the 8th century. It was written in Middle Persian and Manichean scripts.

New Iranian Languages

This section addresses the main literary languages spoken today, many of them now also in the diaspora, especially in Europe and America.

Various forms of Persian (*see* **Persian, Modern**) are spoken throughout Iran, Afghanistan, Tajikistan, and in adjacent areas. It is written in Arabo–Persian, Cyrillic (Tajik), and Hebrew scripts (Judeo–Persian).

Ossetic (ossette) is spoken in Ossetia, in the southern Caucasus, in two main variants, Digoron (the more archaic) and Iron, with subvariants. It is written in the Cyrillic script, except in the south, where Georgian is used.

Kurdish (northern, central, and southern) is spoken in three principal variants in eastern Turkey and Syria, northern Iraq, and western Iran, as well as in surrounding areas. It is written in a modified Persian or standard Turkish Latin script.

Balochi (*see* **Balochi**; or Baluchi, in several dialects) is spoken in eastern Iran and western Pakistan, but also in southern Afghanistan and Central Asia. It is written in the Arabo–Persian script.

Pashto (*see* **Pashto**; several dialect groups) is spoken mainly in Afghanistan and Pakistan. It is written in the Arabo–Persian script.

All the modern languages contain a large number of Arabic words and a smaller number of Turkish words. The easternmost dialects have also borrowed extensively from Indic languages. Differently, Ossetic has borrowed extensively from the neighboring Caucasian languages.

Nonliterary languages and dialects comprise the following:

Northwestern and central Iran: Taleshi on the western shore of the Caspian Sea and Tati dialects from Iranian Azerbaijan through the Central Province and into Gilan; the Caspian dialects (Gilaki in Gilan, Mazanderani, and several dialects on the northern edge of the great salt desert, the Dasht-e Kavir); Gurani (Bajelani) and (including Awromani) in eastern Iraq and western Iran; and Zaza (Dimli) in eastern Turkey; and the Central dialects, comprising a number of more or less interrelated dialects spoken north and south of the Dasht-e Kavir.

Southwestern and southern Iran: Lori (Luri) (in several varieties) and Bakhtiari; Fars dialects; Larestani (in several dialects) in Larestan; dialects in the area from Bandar-e ʻAbbas (Bandari) and Hormoz to Minab and Bashkardi in Bashakerd; and Kumzari on the Musandam peninsula across the Strait of Hormoz (it is unknown whether this is spoken today).

Afghanistan and Central Asia: Parachi and Ormuli (Ormuri) in central Afghanistan and across the border in Pakistan; Ishkashmi and Sanglechi to the west of the Wakhan corridor; Yidgha and Munji/Munjani in eastern Afghanistan and western Pakistan respectively; the Yazghulami-Shughni (Yazgulami-Shugni) group in northern Afghanistan and Central Asia, including Sarikoli in western Xinjiang; Yaghnobi (Yagnobi) in the Yaghnob valley in Tajikistan; and Wakhi in the Wakhan corridor in northeastern Afghanistan.

Genetic Relationships among the Iranian Languages

Only the Persian languages are descended in a more or less direct line. Of the other dialects, the following are more closely related than others: Yidgha–Munji and Bactrian; Yaghnobi and Sogdian; and Wakhi and Khotanese.

Characteristic Phonological Features of Iranian Languages

Phonetic developments differentiating Iranian from its Indo-Iranian ancestor and so also distinguishing it from Old Indic include the following (*see also* **Indo–Iranian**):

1. Indo-Iranian *s* after vowels became Iran. *h* (e.g., OInd. *asura-*, Av. *ahura-* ʻlordʼ).
2. Voiced aspirated stops and affricates lost the aspiration (e.g., OInd. *bhara-*, OIran. *bara-* ʻcarryʼ; OInd. *dhā-* ʻto placeʼ and *dā-* ʻto give,ʼ OIran. both *dā-*; Indo-Iran. **ǰ*^{h} *an-* ʻstrikeʼ > OInd. *han-*, Av., OPers. *jan-*).
3. Indo–European palatal velars (*k̂, ĝ, ĝ*^{h}) and palatalized velar stops (*k*^{y}, *g*^{y}, *g*^{yh}) developed differently in Indic and Iranian (approximate phonetic values: *ś* [ʃ], *ć* [tʃ], *ź* [ʒ], *j́* [dʒ], *š* [ʂ], *č* [tʂ], *ž* [ʐ], *ǰ* [dʐ>]):

IEur.	proto-IIr.	OInd.	proto-Ir.
**k̂, ĝ, ĝ*^{h}	**ć, j́, j́*^{h}	*ś, ǰ, h*	**ć*, **j́*, **j́*
**k, g, g*^{h}	**k*^{y}, *g*^{y}, *g*^{yh} > **č, ǰ, ǰ*^{h}	*č, ǰ, h*	**č*, **ǰ*, **ǰ*

4. The Indo-Iranian unvoiced stops *p, t, k* became spirants *f*, *θ*, *x* before other consonants, including original laryngeals (*H*) (OInd. *cakra-*, Av. *caxra-* ʻwheelʼ; OInd. *trita-*, Av. *θrita-* ʻthirdʼ; OInd. *priya-*, Av. *friia-* ʻdearʼ; Indo-Iran. **ratHa-*, OInd. *ratha-*, Av. *raθa-* ʻchariotʼ; but **pt* → **pt* > Av. *pt*: Av. *hapta* ʻseven,ʼ NPers. *haft*).
5. The geminated dentals *-tt-*, *-dd*$^{(h)}$*-* (*-t*^{s}*t-*, *-d*^{z}d$^{(h)}$-) became *-st-*, *-zd-* (OInd. *vitta-*, Av. *vista-* ʻfoundʼ < *vid-*; OInd. *addhā* ʻtruly,ʼ OPers. *azdā* ʻwell-knownʼ).
6. Indo–Iranian laryngeals remained between vowels but were lost between consonants (Iran. **daHah* ʻgift,ʼ OAv. *daʼō*, spelled *då*; Indo–Iran. **pHtar-* ʻfatherʼ > OAv. *ptar-*, OInd. *pitar-*), although a vowel was inserted (or the vocalized laryngeal kept) in initial groups (YAv., OPers. *pitar-*).

Proto-Iranian in turn split into several distinct dialect groups characterized, among other things, by the developments of the palatal affricates *ć, j́* and the groups *ćw* and *j́w*.

Proto-Ir.	SW-Iran.	Central-Iran.	NE-Ir.
**ć, ǰ́*	*θ, d*	*s, z*	*s, z*
**ćw, ǰ́w*	*s (θ), z*	*sp, zb*	*ś, ź*

To the southwestern group belong the Persian languages and the other languages of southwestern and southern Iran (Bakhtiari and Lori, Fars dialects, Larestani, and Bashkardi). To the northeastern group belong Khotanese and Wakhi. All the others belong to the central group. Examples: Indo–Iran. **daća* '10,' OPers. **daθa*, Av. *dasa*, Khot. *dasau*; Indo–Iran. **aǰ́ʰam* 'I,' OPers. *adam*, Av. *azəm*, Khot. *aysu*; Indo–Iran. **aćwa-* 'horse,' OPers. *asa-*, Median, Av. *aspa-*, Khot. *aśśa-*, Wakhi *yīš*; Indo-Iran. **-ijwā(n)-* 'tongue,' OPers. *hizān-*, MPers. *izbān*, Khot. *biśāa-* /βiźāa-/.

Characteristic Features of Morphology and Points of Syntax

Avestan and Old Persian are still of the Indo–European type (like Greek, Old Indic), with complex morphologies.

In the nominal and pronominal systems, Old and Young Avestan have three numbers (singular, dual, plural), eight cases (nominative, accusative, genitive, dative, instrumental, ablative, locative, vocative); in Old Avestan, the ablative singular has a distinct ending only in the *a*-declension (*-āt̰*), while in Young Avestan, the final *-t̰* is found in all declensions. Old Persian has only six cases and was also losing dual forms.

The Avestan verbal system is based on three stems: present, aorist, and perfect (e.g., *kar-* 'to do': PRES *kərənao-*, AOR *kar-/car-*, PERF *ca-kar-*). All the Indo-Iranian moods are preserved. There is a past participle in *-ta-* (e.g., *kərə-ta-*) and several infinitives. In Old Avestan, the three stems are associated with different aspects: not completed event (present, imperfect, injunctive); completed event (aorist); present result of past event (perfect). In Young Avestan and Old Persian, the aspect system survives mainly in modal forms; in Old Persian, aorist modal forms form a suppletive paradigm with present forms. The Young Avestan present injunctive (present stem with secondary endings) and the Old Persian imperfect (augment plus present stem with secondary endings) became general narrative tense, expressing continuous or punctual actions and events, as well as anteriority, depending on context (YAv. *kərənaom* [*-aom* < *-aw-am*] 'do-1ST.SING,' OPers. *a-kunav-am* 'PAST-do-1ST.SING' = 'I did, (when . . .) I had done.'

An innovation common to Young Avestan and Old Persian is the optative used (with or without augment) to express habitual actions or events in the past (YAv. *apataiiən* < *a-pat-aiy-ant* 'PAST-fall-OPT-3RD.PL' = '(the demons) used to run about'; OPers. *avājanyā* < **ava-a-jan-yā-t* 'down-PAST-strike-OPT-3RD.SING' = 'he used to kill'). This usage continues in several Middle Iranian languages.

Old Persian is in the process of developing a split-ergative verbal system, with an ergativic perfect tense contrasting with the ancient imperfect (*a-kunav-am* 'PAST-*do*-1ST.SING' = 'I did/had done' versus *manā kar-t-am* 'I.GEN-DAT do-PAST.PART-NOM.SING. NEUT' = 'I have done [it]').

The Middle and New Iranian languages exhibit a variety of developments that represent smaller or greater innovations compared with Old Iranian.

Points of Historical Phonology

We may distinguish between languages with conservative phonologies, characterized by the survival of surd stops and affricates after vowels (comparable to Italian in this respect), and languages where these are voiced or dropped (comparable to Spanish and French). Sogdian and Balochi are of the former type, the others at various stages of the latter.

Final syllables are lost to various degrees, e.g., OIran. NOM.SING **aćw-ah* 'horse': Av. *asp-ō*, OPers. *as-a*, Khot. *aśś-ä*, Sogd. *asp-í*, Parth., MPers., NPers. *asp*, Pashto *ās*.

The Indo–Iranian syllabic *r* <r̥> developed differently from *ar* in most languages (Av. <ərə>, OPers. <ar>, MPers. *ir/er, ur*, etc.; e.g., Av. *zərəδaiia-* 'heart,' Parth. *zirδ*, MPers. *dil* [*-rd* > *-l*]; Av. *mərə-ta-* 'dead,' OPers. *mar-ta-*, MPers. *murd*).

All the Old Iranian diphthongs were monophthongized. Fronting and backing of vowels (umlauting) is common in both diachrony and synchrony.

In Chorasmian, Khotanese, and many modern East Iranian languages, the old palatal affricates *č* [tʃ] and *ǰ* [dʒ] became dental affricates *c* [ts] and *j* [dz], while new palatal affricates developed from the velar stops by palatalization.

The groups *xt, ft* became *γt, βt* (with various further developments) in Ossetic and eastern languages (Av. NOM.SING *duxt-a* 'daughter,' Sogd., Chor. *δuγd-á*, Khot. *dūt-a* /δūd-a/, Bactr. <logd-a> = *lugd-a*, Pashto *-lә́* [< **luw* in *tәr-lә́* 'uncle/aunt's daughter']; MPers. *haft* '7,' Sogd. *aβt*, Khot. *hauda*, Pashto *owә́*, Shughni *ūvd*, Wakhi *ɨb*).

Some modern East Iranian languages also have retroflex series, e.g., sibilants *š* [ʃ], *ṣ̌* [ʂ], *ž* [ʒ], *ẓ̌* [ʐ]; affricates *ts, tš, tṣ̌, dz, dž, dẓ̌*, etc. In Pashto, the phonetic realization of the retroflex sibilants defines four dialect areas, going from southwest to northeast: *pašto, paṣ̌to, paxˇto, paxto*.

The development of consonant groups is especially varied:

θr: OPers. *ç*, MPers., etc., *s*; Sogd. (and several mod. dialects) *š*; Khot. *dr-,-r-*; Parth., Bactr. *hr*; Ossetic *rt* (e.g., OIran. NOM.SING **puθr-ah* 'son': OPers. *puç-a*, MPers. *pus*, NPers. *pesar* [changed after *pedar* 'father']; Sogd. *pəš-í*, Khot. *pūr-ä*, Parth. *puhr*, Bactr. <pouro> = *puhr*; Oss. *firt*; OIran. **θrāi̯ah* 'three': MPers. *sē*, NPers. *se*; Sogd. *šē*; Khot. *drai*; Parth. *h(e)ray*, Bactr. <uarēio> = *harēy*, Shughni *aray*; Oss. *ærtæ*, Wakhi *truy*; Munji *x̌iray*).

rs: remains in Ossetic and several eastern languages but commonly becomes *š* in western languages; Pashto and some others have *ṣ̌t* (e.g., **pr̥sa-* 'ask,' Pashto *puṣ̌t-*); the Shughni group has variously *x̌c, ws* (Shughni *pex̌c*, Roshani *paws-*).

rt: frequently remains or becomes *rd* (Wakhi *mərt* 'dead', MPers. *murd*); it becomes retroflex *ṭ* [ʈ] in several eastern languages (e.g., Khot. *muḍa-*), which further becomes retroflex flap *ṛ* [ɽ] (e.g., Pashto *mṛə*), or loses retroflexion and becomes *d* or *g* (Shughni *mūd*, Roshani *mɛwg*).

št: Oss. *st*; Bactr., Pashto *t*; Shughni group *x̌t*; Yidgha-Munji *šk^y*, *šč* (Av. *ašta* 'eight,' OPers. *aštā*, MPers., NPers. *hašt* [with *h-* from *haft* 'seven'], Bactr. <atao> = *ata*, Pashto *atə́*, Shughni *wax̌t*, Yidgha *aščo*).

Points of Historical Morphology and Syntax

Three genders are still found in Sogdian and Khotanese. Two genders are found in Bactrian and Chorasmian and in many modern languages. Persian and many modern languages have no gender.

The dual remains in Sogdian and Chorasmian. The numerative forms used only after numerals in Sogdian and Pashto incorporate old dual forms.

In several Middle and New Iranian languages, marked plural forms are restricted to animate nouns or are used to express individuality.

Definite articles from demonstrative (or relative) pronouns are found in Sogdian, Bactrian, Chorasmian, and Ossetic. In the modern languages, indefiniteness is usually expressed by a suffixed 'one.' Definiteness (topicalization, etc.) is expressed in a variety of ways (which is addressed later, in the section on direct object marking).

The western languages (Middle Persian, Parthian, Bactrian) have reduced the older, six-case system to a two-case system (direct case for the older nominative and accusative[?] and oblique for the other cases), while the eastern languages (Tumshuqese, Khotanese, Sogdian) preserve the Old Persian-type six-case system. In several declensions in Sogdian, in Chorasmian, and in the later stages of Khotanese, the nominative and accusative singular are no longer distinguished (exceptions in the pronouns). Several modern languages preserve the two-case system, sometimes with the addition of one or more local cases. Balochi and other eastern languages have nominative, genitive (possessive), oblique, and a case for the direct/indirect object. Ossetic, to some extent influenced by Caucasian languages, has the largest number of cases, including nominative, genitive, and dative (grammatical cases) and (local cases) allative, ablative, inessive, adessive, equative (expressing language and likeness), and comitative (only Iron).

In Sogdian, a system of light versus heavy stems developed: stems containing at least one long vowel or a diphthong (including *ar, an, am*) attracted the stress, causing loss of short final vowels (e.g., NOM.SING *asp-í* 'horse,' LOC.SING *asp-yá*, versus NOM.SING *mḗθ* 'day,' LOC.SING *mḗθ-ī*; IMPERF.3RD.SING *wan-á* 'did' versus *wḗn* 'saw').

Several Middle and New Iranian languages use affixes to mark the direct object, depending on degree of specificity and definiteness. The most common contrast is unmarked = indefinite versus marked = definite. The markers are of three main types, all older indirect object markers: (1) ending derived from older case ending (usually *-e*, *-i*); (2) from older prepositions meaning 'to' (Manichean Parthian [less commonly Middle Persian] *ō* and Bactrian *abo* from OIran. *abi* 'to([ward])'; Bashkardi *be-* from OIran. *pati* 'to([ward])'); (3) from older prepositions meaning 'from, on account of,' etc. NPers., Balochi *-rā* [from OPers. *rādiy* 'on account of,' MPers. *rāy*, indirect object marker]; Shughni group *a(z)-* [OIran. *hača* 'from']).

In the eastern Middle Iranian languages, as well as in several modern East Iranian languages, the noun has two distinct declensions, one going back to the old vocalic stems, the other to extended *ka*-stems (feminine variously, *kā-, kī-, čī-*, and, analogically, *čā*-stems).

The western languages (Middle Persian, Parthian, Bactrian) continue, to varying extents, the Old Persian use of the relative pronoun as relative particle or *ezafe* (Mid. Pers. *ī*, Parth. *čē*, Bactr. *i-*). In New Persian and related dialects, the use of the ezafe is common. The so-called inverted ezafe found in some languages is the oblique case, e.g., Mazanderani *pēr-e kiya* 'father-OBL house' = 'the father's house.' Possession is of the type 'I have' or 'for me is.'

In verb morphology, innovations include the development of new marked present (continuous) tenses and the restriction of the unmarked old present to, sometimes remarked, modal functions; restructuring of the stem systems into pairs of intransitive-passive and transitive–causative verbs; the restructuring of the past tense systems by the addition of the new

perfect system seen in Old Persian; and, when this becomes the regular past narrative tense, the creation of a new perfect system, using various strategies.

Continuous or progressive tenses are marked already in Sogdian, where a particle meaning 'being' is added to the personal forms (*δār-am-skun* 'hold.PRES-1ST.SING-PROG' = 'I am holding'). In Khotanese, petrified participles meaning 'sitting,' 'standing,' 'going' probably modify the verb in the same sense. In the modern languages, a variety of affixes are found. In Persian and the southwestern dialects, we find prefixes originally meaning 'always' (Class. NPers. *hamē* and *mē-*, NPers. *mī-*; Bakhtiari and Lori *ī-* and *ei-*). The Tati and Caspian dialects use a present participle in **-ende* + copula (e.g., **kun-ende-ī* > *kennī* 'do-PRES.PART-COP.2ND.SING' = 'you are doing, you do'). Several languages use constructions of the type 'be in doing': Gilaki *kār-a āmon dar-a* 'work-CONN come.INF in-COP.3RD.SING' = 'he is coming'; Larestani *a-kerdā-em* 'a-doing I am, I am doing'; NBashk. *a-kerdén-om*, SBashk. *be-kért(én)-īn*; Balochi *raw-ag-ā int* 'a-go-ing he is, he is going'. Some eastern languages use a suffix derived from 'stand' (Yaghnobi *-išt*, e.g., *šáw-om-išt* 'go-1ST.SING-CONT' = 'I am going,' imperf. *a-šáw-i-m-išt* 'PAST-go-IMPERF-1ST.SING.CONT' = 'I was going').

The old present acquires modal functions and can be marked (e.g., NPers. subjunctive *kon-am* 'do-1ST.SING' or *be-kon-am* 'SUBJ.do-1ST.SING' = 'I may do').

The imperfect is still characterized by the augment in Middle Persian (only example *a-ger-īy* 'PAST.do-PASS.3RD.SING' = 'was made,' cf. OPers. *a-kar-iya*), Chorasmian, Sogdian, and Tumshuqese (*a-cchu* 'I went'), and in modern Yaghnobi (*a-kún-i-m* 'PAST-do-IMPERF-1ST.SING' = 'I did'). In Chorasmian and Sogdian, the augment is lost in verbs without original prefix; in verbs with prefix, the augment appears as lengthening of the vowel of the original prefix (e.g., Sogd. *θβar-*, IMPERF *θ-ā-βar-*, Chor. *haβir-*, IMPERF *h-ā-βir-* 'gave,' cf. OPers. *frā-bara* < *fra-a-bara* 'forth-PAST-carry-3RD.SING' = 'he gave') or as *m-* (e.g., Sogd. *m-anγaz-* 'began', Chor. *m-ikk-* 'did', analogically from forms with prefix *ham-*, cf. OPers. *ham-a-taxša-iy* 'PREV-PAST-labor-1ST.SING.MID' = 'I labored'). Several modern languages and dialects have imperfects with a suffix *-i-*, which may continue the Old Iranian preterital optative (e.g., Yaghnobi IMPERF *a-šáw-i-m* 'I was going,' PRES.SUBJ *šáw-om* '(that) I go').

In Middle Persian, Parthian, and Bactrian, the (narrative) past tense system (completed action) is based on the older split ergative seen emerging in Old Persian; Sogdian and Choresmian instead use the auxiliary 'to have' in transitive constructions and 'to be' in intransitive ones (Sogd. *əκ^r-t-u-δār-am* 'do.PAST-ACC.NEUT.SING-have.PRES-1ST.SING' = 'I did, I have done'); and Khotanese has a form with an originally active (possessive) participle plus 'to be' (*dä-t-aimä* 'see-PAST-1ST.SING.MASC' = 'I saw, I have seen' < **dīta' āh ahmi* 'having-seen I am'). As the old original perfect replaced the inherited imperfect (injunctive) as past narrative tense, new strategies were invented to express the perfect. The most common is the use of a past stem extended by *-ag* > *-e*, etc. (NPers. *kárd-am* 'I did', *kard-é-am* 'I have done') or *-ss-* (from the older 'stand'; Larestani *če-d-e* 'go-PAST-PERF.3RD.SING' = 'he has gone,' but *če-ss-em* [*-d-s-* > *-ss-*] 'go-PAST.PERF-1ST.SING' = 'I have gone,' Kumzari *zur-s-e* [*zur-* < *zat-*] 'he has struck').

Punctual aspect is often marked by the prefix *b-* or similar, which also marks modal functions in many modern Iranian languages.

The old modal forms were generally preserved in Middle Iranian: subjunctive, optative, and imperative everywhere; the injunctive (with various modal functions) in Chorasmian, Sogdian, and Khotanese. In the modern languages, old modal forms survive as archaisms (NPers. *zende b-ād* 'alive be.3RD.SING.SUBJ' = 'long live!'), and, instead, modal functions are left unmarked in contrast to the marked present (continuous) or are marked by a prefix, often *b-*.

Directional prefixes are common, more or less closely connected with the verb. For instance, the Late Khotanese adverbs *vā*, *ttā*, *hā* and the Pashto prefixes *rā*, *dər*, *wər* indicate the direction of the action or movement toward first, second, or third person and can also simply substitute for the personal pronouns. Ossetic has a complex system of spatial prefixes indicating the direction of the motion relative to location and speaker.

See also: Avestan; Bactrian; Balochi; Chorasmian; Indo–Aryan Languages; Indo–Iranian; Iran: Scripts, Old Persian, Aramaic, Avestan; Khotanese; Pahlavi; Pashto; Persian, Modern; Persian, Old; Sogdian.

Bibliography

Brill (1958). *Handbuch der Orientalistik. I. Iranistik. IV. 1. Linguistik.* Leiden and Cologne: Brill.

Schmitt R (ed.) (1989). *Compendium Linguarum Iranicarum.* Wiesbaden: Reichert.

Schmitt R (2000). *Die iranischen Sprachen in Geschichte und Gegenwart.* Wiesbaden: Reichert.

Skjærvø P O (2005). 'Iranian languages and their development.' In Yarshater (ed.) vol. 7.

Windfuhr G (ed.) (2005). *The Iranian languages and typology.* London: Curzon.

Yarshater E (ed.) (1982–). *Encyclopaedia Iranica.* London: Routledge and Kegan Paul.

Iraq: Language Situation

B Mahmoodi-Bakhtiari, Center for the Great Islamic Encyclopedia, Tehran, Iran

The Republic of Iraq is home to a variety of different ethnic groups who speak different languages. Of the total 25 374 691 population (2004 census), it is estimated that 75 to 80% of the people are Arabs; 15 to 20% are Kurds; and the remainder, about 5%, is made up of Turkmans, Assyrians, Armenians, and other relatively small groups.

Arabic is the official language of Iraq and is regarded as the mother tongue of more than three-quarters of the population. The written language is a modified version of Classical Arabic, but a diverse set of spoken varieties of Arabic are used. Different dialects of Arabic spoken in Iraq include Mesopotamian Arabic (with almost 12 million speakers), with the alternate names Iraqi Arabic or Furati; North Mesopotamian Arabic, also known as Syro-Mesopotamian Vernacular Arabic, Moslawi, or Mesopotamian Qeltu Arabic; Gulf Arabic, also known as Khaliji, with the dialect Zubair Faau; Judeo-Iraqi Arabic, a greatly endangered dialect, also known as Iraqi Judeo Arabic, Baghdadi Arabic, or Yahudic; and Najdi Arabic, spoken in the south between the rivers up to the Syrian border, with the dialects North and Central Najdi.

Generally, there are two broad groups of spoken Arabic: (a) The spoken Arabic of the Muslims of central and southern Iraq, which can be divided into two subgroups: urban and rural/nomadic. It resembles Gulf Arabic. The Muslim Arabic dialect of Baghdad belongs to this group, and acts as a lingua franca. (b) The spoken Arabic of non-Muslim Arabs, as well as of Muslim Arabs in the north, notably in the province of Mosul. This dialect has marked similarities with some Syrian sedentary dialects.

Iranian languages are the second most commonly used in Iraq. Both Northwestern and Southwestern Iranian languages are used. Among these languages, Kurdish [Kurdi] has the largest community. Kurds have a standard written language and several dialects such as Bajelani, Mukri, Garrusi, Sulaymani, Kolya'I, Zangana, Kermanshahi, Surchi, Herki, and Kurmanji or Northern Kurdish, with the dialects Hakari and Jezire. Other important languages of this group are Gurani, also known as Hawrami or Hawramani, and Zaza [Dimili]. Among the Southwestern Iranian languages, the best known are Lori [Luri] (with the dialects Feyli, Laki, and Bakhtiari), and Persian, which is often spoken along the Iraq-Iran border, and in some Shi'ite holy cities.

Iraq is also interesting for the Aramaic languages spoken in it. The dialects of Neo-Aramaic are spoken in some Christian villages in the north, as well as in central Iraq. Some important members of the Aramaic language family in Iraq are Assyrian Neo-Aramaic or Neo-Syriac, used by the Assyrians, who constitute less than 1 percent of the population and live mainly in the major cities and in the rural areas of northeastern Iraq. Koy Sanjaq Surat is an Aramaic language with 800 to 1000 speakers in the Northern Iraqi town of Koi-Sanjaq and the nearby village of Armota. Chaldean Neo-Aramaic has 100 000 to 120 000 speakers as of 1994.

Iraq, especially its northern part, has been home to diverse linguistic minorities, some of whom were there long before the Arab influx. Technically speaking, Iraq hosts members of Altaic, North Caucasian, and Gypsy languages, in addition to those already noted. The most important Caucasian languages spoken in Iraq are Adyghe (also known as West Circassian or Adygey) and Armenian. Armenian, with about 60 000 speakers, is spoken by an urban community of Armenian descent concentrated in the larger cities.

Of the Altaic group, Azerbaijani and Turkman [Turkmen] are noteworthy. Azerbaijani is spoken mainly in the cities of Kirkuk, Arbil, and Rowanduz, as well as in towns and villages in the southeast from Kirkuk as far as Al Miqdadiyah, Khanaqin, and Mandali, and in several places in the Mosul region. The Turkomans are believed to constitute somewhat less than 2% of the population, and live along the border between the Kurdish and Arab regions. A number of Turkomans also live in the city of Irbil.

Of the minor languages used in Iraq, we should also mention Mandaic, used by the Mandaian people living by the Iranian border with Iraq, and some in Khuzestan province in Iran. The total number of Mandaic speakers is not known. They have a standard written language, though most are fluent in Arabic and Persian.

The only known Gypsy language spoken in Iraq is Domari, with 50 000 speakers as of 1970. This language is also known as Middle Eastern Romani. Nearly all minority language speakers are bilingual in their mother tongue and Arabic.

See also: Arabic; Indo–Iranian; Minorities and Language; Variation in Arabic Languages.

Language Maps (Appendix): Map 102.

Bibliography

Campbell G L (1991). *Compendium of the World's Languages*. London: Routledge.

Comrie B (ed.) (1987). *The World's major languages*. New York: Oxford University Press.

Ethnologue: Languages of the World. 13th ed., (1996). On the WWW at http://www.sil.org/ethnologue.

Fleischman E (1982). 'Proto-Iraqi-Kurdish.' In *Historical and comparative*. Research Papers of the Texas SIL at Dallas. 12. Dallas: Summer Institute of Linguistics. 45–91.

Hoberman R D (1989). *The syntax and semantics of verb morphology in modern Aramaic: a Jewish dialect of Iraqi Kurdistan* (American Oriental Series 69). New Haven: American Oriental Society.

Ingham B (1994). *Najdi Arabic: Central Arabian*. John Benjamins.

Kloss H & McConnell G D (eds) (1984). *Linguistic composition of the nations of the world*. Les Presses de l'Université Laval, Québec; Ruhlen.

Merrit (1987). *A guide to the World's languages, vol. 1: Classification*. Stanford, CA: Stanford University Press.

Sabar Y (1993). 'A folktale and folk songs in the Christian Neo-Aramaic dialect of Tel-Kepe (Northern Iraq).' In Contini R, Pennacchietti F A & Tosco M (eds.) *Semitica. Serta philologica Constantino Tsereteli dicata*. Torino. 289–298.

Ireland, Republic of: Language Situation

C Ó Dochartaigh, Glasgow, UK

This article is reproduced from the previous edition, volume 4, p. 1775, © 1994, Elsevier Ltd.

English and Irish are used in the Republic of Ireland, with English the first language of perhaps 96% of the population, all literate. A dialect of Scots, with no official status, is used by several thousand in a northern county. Irish was adopted as one of the symbols of national identity following independence from the United Kingdom, and a policy of language revival has been pursued mainly through the education system, where Irish is taught to all pupils throughout their schooling.

Although this has succeeded in encouraging large numbers to learn the language, little progress has been made in the area of language use. Irish is returned in census reports as known by over one million, about 30% of the population, but because of self-reporting it is difficult to establish real levels of fluency and literacy. Native speakers of Irish have been recently estimated at around 2%, perhaps 60 000 in all, with about 10 000 living in viable language communities in two western countries. In addition, there are families in English-speaking areas where both parents are fluent Irish learners and where Irish is the everyday language. There are no estimates of the numbers of such families, but a figure of around 2% of the population would not be unreasonable. All Irish speakers are bilingual in English, and code switching is common among native speakers of Irish but not between fluent Irish learners. The Constitution recognizes two official languages, Irish, 'the first,' and English, but there is no statutory provision to formalize their relative status. Most Parliamentary business is transacted through English, with all Acts and most official forms being available in both languages. The main language of the media is English, with an Irish language radio channel. Instruction in most schools and in higher education is through English, but there are a number of Irish-medium schools, both primary and secondary. French, German, Spanish, and Italian are taught as foreign languages in secondary schools.

See also: United Kingdom: Language Situation.

Language Maps (Appendix 1): Map 139.

Irish Lexicography

D P Ó Baoill, Queen's University, Belfast, UK

The first Irish dictionaries were compiled more than a thousand years ago. They consisted mostly of word lists and glosses. It is difficult to evaluate these glossaries until further detailed study of them has been completed.

Very little has been written about the history of Modern Irish (Irish Gaelic) lexicography, and criticism of published dictionaries is seldom based on scientific principles of lexicography. Almost all the main dictionaries are bilingual and are for the most part directed at learners of Irish. This is to be expected, since less than 5% of the population acquire Irish as a first language. Use of these dictionaries depends

Comrie B (ed.) (1987). *The World's major languages*. New York: Oxford University Press.
Ethnologue: Languages of the World. 13th ed., (1996). On the WWW at http://www.sil.org/ethnologue.
Fleischman E (1982). 'Proto-Iraqi-Kurdish.' In *Historical and comparative*. Research Papers of the Texas SIL at Dallas. 12. Dallas: Summer Institute of Linguistics. 45–91.
Hoberman R D (1989). *The syntax and semantics of verb morphology in modern Aramaic: a Jewish dialect of Iraqi Kurdistan* (American Oriental Series 69). New Haven: American Oriental Society.
Ingham B (1994). *Najdi Arabic: Central Arabian*. John Benjamins.
Kloss H & McConnell G D (eds) (1984). *Linguistic composition of the nations of the world*. Les Presses de l'Université Laval, Québec; Ruhlen.
Merrit (1987). *A guide to the World's languages, vol. 1: Classification*. Stanford, CA: Stanford University Press.
Sabar Y (1993). 'A folktale and folk songs in the Christian Neo-Aramaic dialect of Tel-Kepe (Northern Iraq).' In Contini R, Pennacchietti F A & Tosco M (eds.) *Semitica. Serta philologica Constantino Tsereteli dicata*. Torino. 289–298.

Ireland, Republic of: Language Situation

C Ó Dochartaigh, Glasgow, UK

This article is reproduced from the previous edition, volume 4, p. 1775, © 1994, Elsevier Ltd.

English and Irish are used in the Republic of Ireland, with English the first language of perhaps 96% of the population, all literate. A dialect of Scots, with no official status, is used by several thousand in a northern county. Irish was adopted as one of the symbols of national identity following independence from the United Kingdom, and a policy of language revival has been pursued mainly through the education system, where Irish is taught to all pupils throughout their schooling.

Although this has succeeded in encouraging large numbers to learn the language, little progress has been made in the area of language use. Irish is returned in census reports as known by over one million, about 30% of the population, but because of self-reporting it is difficult to establish real levels of fluency and literacy. Native speakers of Irish have been recently estimated at around 2%, perhaps 60 000 in all, with about 10 000 living in viable language communities in two western countries. In addition, there are families in English-speaking areas where both parents are fluent Irish learners and where Irish is the everyday language. There are no estimates of the numbers of such families, but a figure of around 2% of the population would not be unreasonable. All Irish speakers are bilingual in English, and code switching is common among native speakers of Irish but not between fluent Irish learners. The Constitution recognizes two official languages, Irish, 'the first,' and English, but there is no statutory provision to formalize their relative status. Most Parliamentary business is transacted through English, with all Acts and most official forms being available in both languages. The main language of the media is English, with an Irish language radio channel. Instruction in most schools and in higher education is through English, but there are a number of Irish-medium schools, both primary and secondary. French, German, Spanish, and Italian are taught as foreign languages in secondary schools.

See also: United Kingdom: Language Situation.

Language Maps (Appendix 1): Map 139.

Irish Lexicography

D P Ó Baoill, Queen's University, Belfast, UK

The first Irish dictionaries were compiled more than a thousand years ago. They consisted mostly of word lists and glosses. It is difficult to evaluate these glossaries until further detailed study of them has been completed.

Very little has been written about the history of Modern Irish (Irish Gaelic) lexicography, and criticism of published dictionaries is seldom based on scientific principles of lexicography. Almost all the main dictionaries are bilingual and are for the most part directed at learners of Irish. This is to be expected, since less than 5% of the population acquire Irish as a first language. Use of these dictionaries depends

Comrie B (ed.) (1987). *The World's major languages*. New York: Oxford University Press.
Ethnologue: Languages of the World. 13th ed., (1996). On the WWW at http://www.sil.org/ethnologue.
Fleischman E (1982). 'Proto-Iraqi-Kurdish.' In *Historical and comparative*. Research Papers of the Texas SIL at Dallas. 12. Dallas: Summer Institute of Linguistics. 45–91.
Hoberman R D (1989). *The syntax and semantics of verb morphology in modern Aramaic: a Jewish dialect of Iraqi Kurdistan* (American Oriental Series 69). New Haven: American Oriental Society.
Ingham B (1994). *Najdi Arabic: Central Arabian*. John Benjamins.
Kloss H & McConnell G D (eds) (1984). *Linguistic composition of the nations of the world*. Les Presses de l'Université Laval, Québec; Ruhlen.
Merrit (1987). *A guide to the World's languages, vol. 1: Classification*. Stanford, CA: Stanford University Press.
Sabar Y (1993). 'A folktale and folk songs in the Christian Neo-Aramaic dialect of Tel-Kepe (Northern Iraq).' In Contini R, Pennacchietti F A & Tosco M (eds.) *Semitica. Serta philologica Constantino Tsereteli dicata*. Torino. 289–298.

Ireland, Republic of: Language Situation

C Ó Dochartaigh, Glasgow, UK

This article is reproduced from the previous edition, volume 4, p. 1775, © 1994, Elsevier Ltd.

English and Irish are used in the Republic of Ireland, with English the first language of perhaps 96% of the population, all literate. A dialect of Scots, with no official status, is used by several thousand in a northern county. Irish was adopted as one of the symbols of national identity following independence from the United Kingdom, and a policy of language revival has been pursued mainly through the education system, where Irish is taught to all pupils throughout their schooling.

Although this has succeeded in encouraging large numbers to learn the language, little progress has been made in the area of language use. Irish is returned in census reports as known by over one million, about 30% of the population, but because of self-reporting it is difficult to establish real levels of fluency and literacy. Native speakers of Irish have been recently estimated at around 2%, perhaps 60 000 in all, with about 10 000 living in viable language communities in two western countries. In addition, there are families in English-speaking areas where both parents are fluent Irish learners and where Irish is the everyday language. There are no estimates of the numbers of such families, but a figure of around 2% of the population would not be unreasonable. All Irish speakers are bilingual in English, and code switching is common among native speakers of Irish but not between fluent Irish learners. The Constitution recognizes two official languages, Irish, 'the first,' and English, but there is no statutory provision to formalize their relative status. Most Parliamentary business is transacted through English, with all Acts and most official forms being available in both languages. The main language of the media is English, with an Irish language radio channel. Instruction in most schools and in higher education is through English, but there are a number of Irish-medium schools, both primary and secondary. French, German, Spanish, and Italian are taught as foreign languages in secondary schools.

See also: United Kingdom: Language Situation.

Language Maps (Appendix 1): Map 139.

Irish Lexicography

D P Ó Baoill, Queen's University, Belfast, UK

The first Irish dictionaries were compiled more than a thousand years ago. They consisted mostly of word lists and glosses. It is difficult to evaluate these glossaries until further detailed study of them has been completed.

Very little has been written about the history of Modern Irish (Irish Gaelic) lexicography, and criticism of published dictionaries is seldom based on scientific principles of lexicography. Almost all the main dictionaries are bilingual and are for the most part directed at learners of Irish. This is to be expected, since less than 5% of the population acquire Irish as a first language. Use of these dictionaries depends

mostly on the purpose for which they were compiled: to inform native speakers of Irish or to instruct native English speakers. The four main dictionaries dealing with modern Irish usage were all published in the 20th century, namely, Fr. Patrick Dinneen's *Irish-English dictionary* (1927), Fr. L. McKenna's *English-Irish dictionary* (1935), Tomás de Bhaldraithe's *English-Irish dictionary* (1959), and Niall Ó Dónaill's *Irish-English dictionary* (1977).

Irish speakers owe the greatest debt to Dineen's work (1927). Its success and popularity has been fostered over generations, as it was the dictionary used to teach Modern Irish to several generations of Irish speakers and learners. It helped standardize Irish grammar and idiom and regulate spelling conventions. De Bhaldraithe's dictionary (1959) was the first to include the new formulated standard spelling and grammar, published in 1958 under the title *Graiméar na Gaeilge: an caighdeán oifigiúil*. All subsequent publications and dictionaries have been written using this standard or some slightly emended form.

The 1959 and 1977 dictionaries were compiled in the Department of Education. Due to the many cultural and linguistic changes undergone by both the Irish and English languages, vocabularies in both languages have undergone and are undergoing major changes. There is, therefore, a need for continuous updating and review.

The *Dictionary of the Irish language*, which was based mainly on Old Irish (600–900 A.D.) and Middle Irish (900–1200 A.D.) materials, was published over a 150-year period, the last fascicle appearing in 1976. A decision was then taken to continue with a similar lexicographical project on Modern Irish, covering the period 1600–1900. This was indeed an urgent necessity. The aim was to produce a monolingual historical dictionary, using the standard modern forms as headwords. A decision was taken in 1984 that the dictionary should appear in the form of fascicles, and the drafting of letter *A* began. Tomás de Bhaldraithe resigned his post as Chief Editor in 1994. Subsequently, a review of the procedures being implemented by the dictionary staff concluded that its objectives needed to be redefined. For the next decade or so, the main objective was to produce a computerized dictionary archive for Modern Irish. This work has now been made available electronically (2004). The work on the dictionary itself remains incomplete.

Phase 2 of the new English-Irish dictionary is due to begin in 2005. It is expected that it will take three to five years to complete. The dictionary will contain some 70 000 headwords and will be corpus-based. The Irish language component in the dictionary will include the three principal varieties spoken in Ireland. The English language in the dictionary will be the language of Ireland, both Northern Ireland and the Republic. While foregrounding Irish English, it will also include standard British and American English and well-known words and phrases from other varieties. An electronic version of the dictionary will also be made available.

See also: Ireland, Republic of: Language Situation; Language Policies: Policies on Language in Europe; Lexicography: Overview; United Kingdom: Language Situation.

Bibliography

Daltún S (1983). 'Scéal Rannóg an Aistriúcháin.' *Teangeolas 17(Winter),* 12–26.

DeBhaldraithe T (1978). 'An Foclóir Nua.' *Feasta (August).* 5–7.

De Bhaldraithe T (1980). 'Foclóirí agus foclóireacht na Gaeilge.' *Maynooth Review 6(1),* 3–15.

De Bhaldraithe T (1990). 'Risteard Pluincéad (fl. 1662) – a neglected pioneer Irish lexicographer.' *Studia Celtica Japonica 3,* 1–8.

De Bhaldraithe T (2004). 'Irish dictionaries.' In *Corpas na Gaeilge/The Irish language corpus.* Dublin: Royal Irish Academy.

Ó Baoill D P & Ó Riagáin P (1990). 'Reform of the orthography, grammar and vocabulary of Irish.' In Fodor I & Hagège C (eds.) *Language reform: history and future V.* Hamburg:. 173–195.

Ó Cuív B (ed.) (1969a). *A view of the Irish language.* Baile Átha Cliath: Oifig an tSoláthair.

Ó Cuív B (1969b). 'Irish in the modern world.' In Ó Cuív B (ed.) (1969a). *A view of the Irish language.* Baile Átha Cliath: Oifig an tSoláthair. 122–132.

Ó Dónaill N (1951). *Forbairt na Gaeilge.* BÁC.

Ó hUiginn R (ed.) (2001). *'Ceist na Teanga.' Léachtaí Cholm Cille XXXI.* Maigh Nuad: An Sagart.

Ó Muirithe D (1996). *The words we use.* Dublin: Four Courts Press.

Ó Riagáin P (ed.) (1985). *International journal of the sociology of language: language planning in Ireland.* Amsterdam: Mouton de Gruyter.

Royal Irish Academy (2004). *Corpas na Gaeilge 1600–1882/The Irish language corpus.* Dublin: Royal Irish Academy.

Irony

S Attardo, Youngstown State University, Youngstown, OH, USA

The term 'irony' is commonly used to describe both a linguistic phenomenon (verbal irony) and other phenomena including 'situational' irony (i.e., irony of facts and things dissociated from their linguistic expression; Shelley, 2001) such as a fire-station burning to the ground, various more-or-less philosophical ideas (Socratic irony, Romantic irony, Postmodern irony), and even a type of religious experience (Kierkegaard, 1966). While there may be connections between situational and verbal irony, it does not appear that literary and religious uses can be fruitfully explained in terms of linguistic irony. This treatment will be limited to verbal irony.

Other definitional problems include the purported distinction between irony and sarcasm. While some have argued that the two can be distinguished (for example, irony can be involuntary, while sarcasm cannot be so), others maintain that no clear boundary exists. A further problem is presented by the fact that in some varieties of English, the term *irony* is undergoing semantic change and is assuming the meaning of an unpleasant surprise, while the semantic space previously occupied by *irony* is taken up by the term *sarcasm*.

The word *irony* goes back to the Greek *eironeia* (pretense, dissimulation) as does the history of its definition and analysis. Irony is seen as a trope (i.e., a figure of speech) in ancient rhetorics and this analysis has remained essentially unchallenged until recently. In the traditional definition irony is seen as saying something to mean the opposite of what is said. This definition is demonstrably incorrect, as a speaker may be ironical but not mean the opposite of what he/she says; cf. *It seems to be a little windy* (uttered in the middle of a violent storm), in which the speaker is saying less than what is meant. Similarly, overstatements and hyperbole may be ironical (Kreuz and Roberts, 1995).

A recent and fruitful restatement of the irony-as-trope theory has been presented by Paul Grice who sees irony as an implicature, i.e., as a deliberate flouting of one of the maxims of the principle of cooperation. Relatedly, speech-act approaches to irony see it as an insincere speech act. Initially, Grice's approach saw irony as a violation of the maxim of quality (i.e., the statement of an untruth) but this claim has been refuted, as seen above. Broadening the definition to, for example, 'saying something while meaning something else,' runs the risk of obliterating the difference between irony and other forms of figurative or indirect speech. However, this loss of distinction may be a positive aspect of the definition, as has been recently argued (Kreuz, 2000, Attardo, 2002).

While the idea of 'oppositeness' in irony is problematic, approaches to irony as negation have been presented (Giora, 1995), who sees irony as 'indirect' (i.e., inexplicit; cf. Utsumi, 2000) negation; related ideas are that of contrast (Colston, 2002) and inappropriateness (Attardo, 2000).

A very influential approach to irony is the mention theory (Sperber and Wilson, 1981), which claims that an utterance is ironical if it is recognized as the echoic mention of another utterance by a more or less clearly identified other speaker. Furthermore, the ironical statement must be critical of the echoed utterance (cf. Grice, 1989: 53–54). Similar theories based on the ideas of 'pretense' and 'reminder' have been presented as well. Criticism of the mention theory notes that not all irony seems to be interpretable as the echo of someone's words, or that if the definition of mention is allowed to encompass any possible mention it becomes vacuous (since any sentence is potentially the mention of another sentence). Furthermore, there exists an admittedly rarer, non-negative, praising irony, called *asteism* (Fontanier, 1968: 150). An example of asteism might be a colleague describing Chomsky's *Aspects of the theory of syntax* as a 'moderately influential' book in linguistics. Other approaches to irony include the 'tinge' theory, which sees irony as blending the two meanings (the stated and the implied ones) with the effect of attenuating the ironical one (Colston, 1997).

All the theories of irony mentioned so far share the idea that the processing of irony is a two-step process in which one sense (usually assumed to be the literal meaning) of the utterance is accessed and then a second sense of the utterance is discovered (usually under contextual pressure). Thus, for example, in a Gricean account of irony as implicature, the hearer of an utterance such as *That was smart* (uttered as a description of clumsy behavior, such as spilling one's wine upon someone's clothing) will first process the utterance as meaning literally roughly 'This behavior was consonant with how smart people behave' and then will discard this interpretation in favor of the implicature that the speaker means that the behavior was *not* consonant with how smart people behave. This account has been challenged recently by 'direct access' theories.

The direct access theories claim that the hearer does not process the literal meaning of an ironical

utterance first and only later accesses the figurative (ironical) meaning. Rather, they claim that the literal meaning is either not accessed at all or only later. Direct access interpretations of irony are squarely at odds with the traditional interpretation of irony as an implicature. Some results in psycholinguistics have been seen as supporting this view (Gibbs, 1994). The mention theory of irony was commonly interpreted as a direct access theory, but recent work (Yus, 2003) seems to indicate that it too can be interpreted as a two-step process. Other researchers (e.g., Dews and Winner, 1999) have presented contrasting views which support the two-step approach, although not always the claim that the literal meaning is processed first: claims that interpretations are accessed in order of saliency (Giora, 2003) or in parallel have been put forth.

Psycholinguistic studies of irony have focused on children's acquisition of irony (Winner, 1988), progressively lowering the age at which children understand irony to under ten years old; on the neurobiology of the processing of irony (McDonald, 2000), emphasizing the role of the right hemisphere alongside the left one (in which most language processing takes place); and on the order of activation of the various meanings in the ironical text. A significant issue is the degree and nature of the assumptions that the hearer and speaker must share for irony to be understood; this can be summed up as the 'theory of mind' that the speakers have. In particular, irony involves metarepresentations (Bara *et al.*, 1997, Curcó, 2000).

Considerable attention has been paid to the optional markers of irony, i.e., primarily intonational and kinesic indications of the speaker's ironical intent. While several phonological and other features have been considered 'markers' of irony, it appears that none of these features is exclusively a marker of irony. Reviews of markers include phonological (e.g., intonation), graphic (e.g., italics, punctuation), morphological (e.g., quotatives), kinesic (e.g., winking), and contextual clues (Haiman, 1998).

Recently, the social and situational context of irony as well as its pragmatic ends have begun being investigated in sociolinguistics and discourse/conversation analysis as well as in psycholinguistics. Work on the social functions of irony has found a broad range of functions, including in- and out-group definition, evaluation, aggression, politeness, verbal play, and many others (e.g., Clift, 1999; Anolli *et al.*, 2002; Gibbs and Colston, 2002; Kotthoff, 2003). It is likely that this list may be open-ended.

The relationship between irony and humor remains underexplored, despite their obvious connections, although some studies are beginning to address the interplay of irony and other forms of implicature, such as indirectness, and metaphoricity. Finally, it is worth noting that dialogic approaches to language (e.g., Ducrot, 1984) see irony as a prime example of the co-presence of different 'voices' in the text (*see* **Literary Pragmatics**), in ways that avoid the technical problems highlighted in the mention theories.

See also: Figures of Speech; Implicature; Negation; Pragmatics: Overview; Relevance Theory; Speech Acts.

Bibliography

Anolli L, Ciceri R & Riva G (eds.) (2002). *Say not to Say: New Perspectives on miscommunication.* Amsterdam: IOS Press.

Anolli L, Infantino M G & Ciceri R (2002). '"You're a real genius!": irony as a miscommunication design.' In Anolli, Ciceri & Riva (eds.). 135–157.

Attardo S (2000). 'Irony as relevant inappropriateness.' *Journal of Pragmatics 32(6),* 793–826.

Attardo S (2002). 'Humor and irony in interaction: from mode adoption to failure of detection.' In Anolli, Ciceri & Riva (eds.).

Bara B, Maurizio T & Zettin M (1997). 'Neuropragmatics: neuropsychological constraints on formal theories of dialogue.' *Brain and Language 59,* 7–49.

Booth W (1974). *A rhetoric of irony.* Chicago: University of Chicago Press.

Clift R (1999). 'Irony in conversation.' *Language in Society 28,* 523–553.

Colston H L (1997). 'Salting a wound or sugaring a pill: the pragmatic function of ironic criticism' *Discourse Processes 23,* 25–45.

Colston H L (2002). 'Contrast and assimilation in verbal irony.' *Journal of Pragmatics 34(2),* 111–142.

Curcó C (2000). 'Irony: negation, echo and metarepresentation.' *Lingua, 110,* 257–280.

Dews S & Winner E (1999). 'Obligatory processing of literal and non-literal meanings in verbal irony.' *Journal of Pragmatics 31(12),* 1579–1599.

Ducrot O (1984). *Le dire et le dit.* Paris: Editions de Minuit.

Fontanier P (1968). *Les figures du discours.* Paris: Flammarion. Originally published as two volumes in 1821 and 1827.

Gibbs R W (1994). *The poetics of mind: figurative thought, language, and understanding.* Cambridge/New York: Cambridge University Press.

Gibbs R W & Colston H L (2002). 'The risks and rewards of ironic communication.' In Anolli, Ciceri & Riva (eds.). 181–194.

Giora R (1995). 'On irony and negation.' *Discourse Processes 19,* 239–264.

Giora R (2003). *On our mind.* Oxford: Oxford University Press.

Haiman J (1998). *Talk is cheap: sarcasm, alienation, and the evolution of language.* Oxford/New York: Oxford University Press.

Katz A N (ed.) (2000). 'The Uses and Processing of Irony and Sarcasm.' Special issue of *Metaphor and Symbol 15(1/2)*.
Kierkegaard S (1966). *The concept of irony, with constant reference to Socrates*. Capel L M (trans.). London: Collins.
Kotthoff H (2003). 'Responding to irony in different contexts: on cognition in conversation.' *Journal of Pragmatics 35(9),* 1387–1411.
Kreuz R J & Roberts R M (1995). 'Two cues for verbal irony: hyperbole and the ironic tone of voice.' *Metaphor and Symbolic Activity 10(1),* 21–31.
McDonald S (2000). 'Neuropsychological studies of sarcasm.' *Metaphor and Symbol 15(1/2),* 85–98.
Shelley C (2001). 'The bicoherence theory of situational irony.' *Cognitive Science 25,* 775–818.
Sperber D & Wilson D (1981). 'Irony and the use mention distinction.' In Cole P (ed.) *Radical Pragmatics*. New York/London: Academic Press. 295–318.
Toplak M & Katz A N (2000). 'On the uses of sarcastic irony.' *Journal of Pragmatics 32(10),* 1467–1488.
Utsumi A (2000). 'Verbal irony as implicit display of ironic environment: distinguishing ironic utterances from nonirony.' *Journal of Pragmatics 32(12),* 1777–1806.
Winner E (1988). *'The point of words: children's understanding of metaphor and irony'*. Cambridge, MA: Harvard University Press.
Yus F (2003). 'Humor and the search for relevance.' *Journal of Pragmatics 35(9),* 1295–1331.

Irony: Stylistic Approaches

J Boutonnet, University of Wolverhampton, Wolverhampton, UK

Most books or articles on the topic start with a warning that the concept of irony is elusive (Muecke, 1970) and that it takes many forms. Myers (1981: 410) said that "in irony we inherit both a device and a concept." Nash (1985) alerted us to the lack of consistency in usage even among literary critics. Although dictionaries offer definitions and exemplifications, they point to the complexity of the phenomenon rather than enlighten the reader. The word has both positive and negative connotations today, a mark of its complex history and range of functions in discourse. The word does not appear in the English language until the 16th century and is only commonly used by speakers from the early 18th century onwards. Clearly the phenomenon existed before it was named, even prior to the Greek word from which the English word derives. The word *eironeia* to describe the Greek philosopher Socrates's treatment of his conversational opponents is first recorded in Plato's *Republic*. Socratic irony today refers to a discourse strategy whereby the speaker pretends he or she is learning from an interlocutor whilst trying all the while to uncover the flaws in that person's argument (Nilsen and Nilsen, 2000).

Conveniently, a distinction is often made between situational irony and verbal irony. In the first we have an observer who notices a state of affairs or event which is in some way incongruous. The second is a rhetorical device. We have an ironist who produces a linguistic message whose meaning is other than the literal meaning, in many cases understood to be the opposite of the literal meaning conveyed. There is a discrepancy between the words and what the speaker means by these words, an opposition between the surface and the underlying meaning. This is also called traditional irony in most writings and is often attributed to Aristotle (Barbe, 1995).

An oppositional model of verbal irony holds that the speaker says the opposite of what he or she means. In this account, irony is unveiled by a process of substitution of lexical items or propositions. In the case of stable irony (Booth, 1974), where the intended meaning is fairly clear, as in "What a lovely day!" uttered on a day when the weather is anything but pleasant, meaning may be retrieved by word substitution (e.g., 'lovely' for 'awful'), an example of lexical-antonymy. Oppositeness is at the heart of Raskin's semantic theory of humor (STH) (Raskin, 1985) where he discussed oppositeness in terms of local antonymy, two opposite meanings arising from a particular discourse. At the propositional level, consider Myers's (1981: 411) example "I always wanted to spend the summer in Detroit," uttered by a speaker who does not want to spend the summer in Detroit. Here we have an example of negation at the propositional level. Some linguists argue that examples like the one above constitute an instance of criticism coupled with a complaint, also known as nonce irony. Common irony refers to expressions which have become stock phrases in a language, such as "that's a likely story!" (Barbe, 1995: 18–22).

When considering verbal irony, a number of assumptions need clarification. One often-debated issue is that of intentionality and sincerity of the

Katz A N (ed.) (2000). 'The Uses and Processing of Irony and Sarcasm.' Special issue of *Metaphor and Symbol 15(1/2)*.
Kierkegaard S (1966). *The concept of irony, with constant reference to Socrates*. Capel L M (trans.). London: Collins.
Kotthoff H (2003). 'Responding to irony in different contexts: on cognition in conversation.' *Journal of Pragmatics 35(9),* 1387–1411.
Kreuz R J & Roberts R M (1995). 'Two cues for verbal irony: hyperbole and the ironic tone of voice.' *Metaphor and Symbolic Activity 10(1),* 21–31.
McDonald S (2000). 'Neuropsychological studies of sarcasm.' *Metaphor and Symbol 15(1/2),* 85–98.
Shelley C (2001). 'The bicoherence theory of situational irony.' *Cognitive Science 25,* 775–818.
Sperber D & Wilson D (1981). 'Irony and the use mention distinction.' In Cole P (ed.) *Radical Pragmatics*. New York/London: Academic Press. 295–318.
Toplak M & Katz A N (2000). 'On the uses of sarcastic irony.' *Journal of Pragmatics 32(10),* 1467–1488.
Utsumi A (2000). 'Verbal irony as implicit display of ironic environment: distinguishing ironic utterances from nonirony.' *Journal of Pragmatics 32(12),* 1777–1806.
Winner E (1988). *'The point of words: children's understanding of metaphor and irony'*. Cambridge, MA: Harvard University Press.
Yus F (2003). 'Humor and the search for relevance.' *Journal of Pragmatics 35(9),* 1295–1331.

Irony: Stylistic Approaches

J Boutonnet, University of Wolverhampton, Wolverhampton, UK

Most books or articles on the topic start with a warning that the concept of irony is elusive (Muecke, 1970) and that it takes many forms. Myers (1981: 410) said that "in irony we inherit both a device and a concept." Nash (1985) alerted us to the lack of consistency in usage even among literary critics. Although dictionaries offer definitions and exemplifications, they point to the complexity of the phenomenon rather than enlighten the reader. The word has both positive and negative connotations today, a mark of its complex history and range of functions in discourse. The word does not appear in the English language until the 16th century and is only commonly used by speakers from the early 18th century onwards. Clearly the phenomenon existed before it was named, even prior to the Greek word from which the English word derives. The word *eironeia* to describe the Greek philosopher Socrates's treatment of his conversational opponents is first recorded in Plato's *Republic*. Socratic irony today refers to a discourse strategy whereby the speaker pretends he or she is learning from an interlocutor whilst trying all the while to uncover the flaws in that person's argument (Nilsen and Nilsen, 2000).

Conveniently, a distinction is often made between situational irony and verbal irony. In the first we have an observer who notices a state of affairs or event which is in some way incongruous. The second is a rhetorical device. We have an ironist who produces a linguistic message whose meaning is other than the literal meaning, in many cases understood to be the opposite of the literal meaning conveyed. There is a discrepancy between the words and what the speaker means by these words, an opposition between the surface and the underlying meaning. This is also called traditional irony in most writings and is often attributed to Aristotle (Barbe, 1995).

An oppositional model of verbal irony holds that the speaker says the opposite of what he or she means. In this account, irony is unveiled by a process of substitution of lexical items or propositions. In the case of stable irony (Booth, 1974), where the intended meaning is fairly clear, as in "What a lovely day!" uttered on a day when the weather is anything but pleasant, meaning may be retrieved by word substitution (e.g., 'lovely' for 'awful'), an example of lexical-antonymy. Oppositeness is at the heart of Raskin's semantic theory of humor (STH) (Raskin, 1985) where he discussed oppositeness in terms of local antonymy, two opposite meanings arising from a particular discourse. At the propositional level, consider Myers's (1981: 411) example "I always wanted to spend the summer in Detroit," uttered by a speaker who does not want to spend the summer in Detroit. Here we have an example of negation at the propositional level. Some linguists argue that examples like the one above constitute an instance of criticism coupled with a complaint, also known as nonce irony. Common irony refers to expressions which have become stock phrases in a language, such as "that's a likely story!" (Barbe, 1995: 18–22).

When considering verbal irony, a number of assumptions need clarification. One often-debated issue is that of intentionality and sincerity of the

ironist. Here the focus is on the production of an ironic text. This approach has been popular in stylistic analyses of literary texts, concerned with identifying devices employed by writers. The rhetorician Quintilian claimed that the ironist 'intends' to convey something which is other than what he or she says, thus effectively being insincere. Albeit not universally held, this view still finds many supporters today. Interestingly, the anti-intentionalist position is refuted on the basis that if the text can be read ironically, it is read "as if it were intended to be ironical" (Muecke, 1978: 364), highlighting the importance of the speech community's conventions of usage over the individual speaker's intentions. However, Gibbs *et al.* (1995) provided convincing evidence that intention is not necessary for irony to occur. They show that hearers are perfectly able to recognize the ironic meaning created by a sentence whilst knowing at the same time that the speaker's intention was not to convey irony. It is true that instances of dramatic irony rely on this, as is the case when readers of a book, or an audience at a play, know something that the characters themselves do not know. In the world of theatre, the *ingénu* who utters sincerely the words of the playwright is a good example of unintentional irony. The speaker is separate from the ironist, the animator of the utterance distinct from the author of the utterance (Goffman, 1979).

Haiman (1990: 188) explained that the distinction between sarcasm and irony is rooted in intention. Whilst sarcasm is intentional, it is not a necessary condition for irony. "To be ironic, a speaker need not be aware that his words are 'false' – it is sufficient that his interlocutors or his audience be aware of this." Sarcasm is often categorized as a form of irony and some researchers use the term interchangeably (Attardo *et al.*, 2003). However, for some, it is a type of indirect criticism which is perhaps more personal and more blunt, and where the speaker's intentions to criticize are more obvious. Barbe (1995: 28) distinguished between the two by considering whose face is more threatened by the utterance. In the case of sarcasm she claims the speaker is compromising him- or herself, and thus a sarcastic comment constitutes a face-threatening act for the speaker.

As for the notion of double significance, and more precisely the clash between the literal and implied meaning, it is clear that the meanings are not necessarily in a relation of 'opposition' in the strict sense, as is often presumed. When making use of hyperbole it is possible to speak the truth whilst conveying an additional meaning. Myers's example of the comment "I love people who signal," uttered by a passenger after the driver attempts a left turn without signaling (Myers, 1977: 172), is a good case in point. Here the speaker is not saying she does not like people who signal, a negation of the declaration, neither is she being sincere about loving all people who signal. As well as exaggeration, understatement may be used ironically, effectively flouting Grice's maxim of quantity – which suggests that your contribution must be as informative as is required, but no more than is required, for the purposes of the current exchange (Levinson, 1983). The ironist is speaking the truth, but implies, by omitting to give as much information as is required, that there is a conversational implicature to be inferred by the hearer (*see* **Maxims and Flouting**). Ironical quotations and interjections do not fit into the traditional oppositional model of irony either.

Irony can make use of metaphors (*see* **Metaphor: Stylistic Approaches**). Overusing metaphors, placing unfitting metaphors, and juxtaposing unlikely metaphors are useful devices for the production of irony. In the case of explicit irony (as in a sentence starting with, for example, 'ironically' or 'isn't it ironic') the speaker brings to the recipient's attention an incongruous occurrence, communicating the nonverbal irony of fate. But in most instances of irony, which are implicit, inference is the key to uncovering the 'true' meaning. However, as with all instances of irony, interpretation is idiosyncratic and it is not certain that meaning will be recovered by the hearer. This has led many researchers to shift their focus away from the ironist to the recipient of the message (*see* **Irony**) (Kaufer, 1981; Giora, 1995; Giora and O'Fein, 1999a, 1999b).

Building on Sperber and Wilson's echoic interpretation model (Sperber and Wilson, 1981; Wilson and Sperber, 1992) and the pretense and theater models (Clark and Gerrig, 1984; Haiman, 1990), Clift (1999) concluded that in naturally occurring conversations, irony is not necessarily located in the utterance itself, thus giving credence to the view that irony cannot reliably be identified on the basis of linguistic description alone. Irony emerges from dynamic conversational activity and is exploited by creative speakers as part of their turn-taking involvement. For Clift, irony may emerge from a flouting of a conversational expectation, a tension between 'slots' and the 'items' that fill them (Clift, 1999: 547). Irony in natural conversation serves an important discourse and social goal and is used in the management of interpersonal relationships (Norrick, 1993). It can be used to criticize whilst not losing face and at the same time minimize the threat to the face of the listener, thus adhering to rules of politeness. It can be used to display wit and solidarity and to affirm bonds between conversationalists too (affiliative vs. nonaffiliative purposes).

In pragmatics it is postulated that recognition of irony lies in the markers used to direct the recipient of the message to the right interpretation. Contextual clues can appear at all levels of language: word, clause, utterance, and discourse. In speech, pointers include kinesic (ranging from gestures to nods and smiles) and phonic (tone of voice, stress, pitch, etc.) markers. Another pointer to irony may come from the contribution being inappropriate in relation to the context or the cotext, even when on the surface the contribution seems relevant. Contradictory clauses are also a trigger. One indicator is the use of register clashes. During an informal conversation one might shift from an informal style to a formal style for ironic effect, for example, or insert phraseology which belongs to another field of discourse. An unusual collocation may alert the speaker (or reader) to the existence of an ambiguity. Jane Austen's opening sentence in *Pride and Prejudice*,

> "It is a truth universally acknowledged that a single man in possession of a good fortune must be in want of a wife,"

combines overstatement ('universally,' 'must be') with a contradiction ('good fortune' / 'in want of') to indicate that the sentence must not be taken at face value. The choice of lexis points to an incongruity and alerts the reader to another intended meaning. Markers, however, are pointers and do not constitute irony *per se*.

Although verbal irony does not always imply a script opposition or negation of the literal message, it allows for a double perspective to be conveyed. This may explain why much irony falls into the realm of humor (*see* **Humor in Language**), as the perception of a contrast is at the heart of Incongruity Theories and Surprise Theories (Goldstein and McGhee, 1972). Attardo (1994) placed the word 'irony' within the semantic field of humor, and claimed that irony and humor overlap significantly (Attardo, 2001). Others have argued that irony and humor share the same mechanisms (Giora, 1995). Irony is used to express an evaluative judgment, to convey an attitude about the interlocutor, an event, or more generally about some aspect of one's universe of belief (Martin, 1992) with the possible outcome that it can have the perlocutionary effect of being humorous, a desirable quality for most. For the recipient, satisfaction is gained from the process of decoding the multiple readings. In literature, techniques used by writers which signal that humor is present also enhance the reader's experience. Triezenberg (2004) explained that, for example, as a device, repetition with skillful variation is a potent humor enhancer. She also drew attention to the importance of cultural factors and shared background knowledge for ensuring that humor is decoded and fully appreciated. In order to achieve a "secret communion" (Leech and Short, 1981) between author and reader, a complex range of linguistic and nonlinguistic factors needs to be combined.

The general theory of verbal humor (*see* **Humor: Stylistic Approaches**) was developed to describe the conditions necessary for a short text to be considered humorous, and aims to provide an explanation of the cognitive processes involved in the processing of humorous texts (Attardo, 2001, 2002). From considering the structure of the joke, the theory has been extended to attempt to analyze humorous literary texts. Although some attest to its literary stylistics value, others argue that it is hitherto not adequate to account for the aesthetic qualities of literary writing (Triezenberg, 2004). Nevertheless, irony as well as humor is a linguistic resource which speakers and writers employ skillfully and acquire as part of their linguistic competence, and recognize in others' performance. No absolute agreement can ever be reached about incidences of irony and their interpretation, as usage as well as interpretations are rooted in the social and linguistic conventions of a particular speech community at a specific point in time. To the extent that any instance of language use can be analyzed using rigorous methods of linguistic and stylistic analysis, it is possible to say something about the mechanisms employed in the use of irony, even if the concept itself is constantly evolving.

See also: Humor in Language; Humor: Stylistic Approaches; Irony; Maxims and Flouting; Metaphor: Stylistic Approaches.

Bibliography

Attardo S (1994). *Linguistic theories of humor.* Berlin: Mouton de Gruyter.

Attardo S (2001). *Humorous texts: a semantic and pragmatic analysis.* Berlin: Mouton de Gruyter.

Attardo S (2002). 'Cognitive stylistics of humorous texts.' In Semino E & Culpeper J (eds.) *Cognitive stylistics: language and cognition in text analysis.* Amsterdam: John Benjamins. 231–250.

Attardo S, Eisterhold J, Hay J & Poggi I (2003). 'Multimodal markers of irony and sarcasm.' *International Journal of Humor Research 16(2),* 243–260.

Barbe K (1995). *Irony in context.* Amsterdam: John Benjamins.

Booth W (1974). *A rhetoric of irony.* Chicago: University of Chicago Press.

Clark H H & Gerrig R J (1984). 'On the pretense theory of irony.' *Journal of Experimental Psychology: General 113,* 121–126.

Clift R (1999). 'Irony in conversation.' *Language in Society 28(4),* 523–553.
Gibbs R W, O'Brien J E & Doolittle S (1995). 'Inferring meanings that are not intended: speakers' intentions and irony comprehension.' *Discourse Processes 20,* 187–203.
Giora R (1995). 'On irony and negation.' *Discourse Processes 19,* 239–264.
Giora R & O'Fein O (1999a). 'Irony: context and salience.' *Metaphor and Symbol 14(4),* 241–257.
Giora R & O'Fein O (1999b). 'Irony comprehension: the graded salience hypothesis.' *International Journal of Humor Research 12(4),* 425–436.
Goffman E (1979). 'Footing.' *Semiotica 25,* 1–29.
Goldstein J H & McGhee P (1972). *The psychology of humor.* New York: Academic Press.
Haiman J (1990). 'Sarcasm as theater.' *Cognitive Linguistics 1(2),* 181–205.
Kaufer D S (1981). 'Understanding ironic communication.' *Journal of Pragmatics 5,* 495–510.
Leech G & Short M H (1981). *Style in fiction.* London: Longman.
Levinson S C (1983). *Pragmatics.* Cambridge: Cambridge University Press.
Martin R (1992). 'Irony and universe of belief.' *Lingua 87,* 77–90.
Muecke D C (1970). *Irony: the critical idiom.* London: Methuen.
Muecke D C (1978). 'Irony markers.' *Poetics 7,* 363–375.
Myers R A (1977). 'Towards a definition of irony.' In Fasold R W & Shuy R W (eds.) *Studies in language variation.* Washington, DC: Georgetown University Press. 171–183.
Myers R A (1981). 'The function of irony in discourse.' *Text 1(4),* 407–423.
Nash W (1985). *The language of humour: style and technique in comic discourse.* London: Longman.
Nilsen A & Nilsen D (2000). *Encyclopedia of humor and comedy.* Westport, CT: Greenwood Press.
Norrick N (1993). *Conversational joking.* Bloomington: Indiana University Press.
Raskin V (1985). *The semantic mechanisms of humor.* Dordrecht: D. Reidel.
Sperber D & Wilson D (1981). 'Irony and the use–mention distinction.' In Cole P (ed.) *Radical pragmatics.* New York: Academic Press. 295–318.
Triezenberg K (2004). 'Humor enhancers in the study of humorous literature.' *International Journal of Humor Research 17(4),* 411–418.
Wilson D & Sperber D (1992). 'On verbal irony.' *Lingua 87,* 53–76.

Iroquoian Languages

M Mithun, University of California, Santa Barbara, CA, USA

The Iroquoian Family

The Iroquoian languages are indigenous to southeastern and northeastern North America. The family consists of two major branches: Southern Iroquoian and Northern Iroquoian.

Southern Iroquoian is represented by just one language, Cherokee, spoken today primarily in Oklahoma and western North Carolina. There are clear dialect differences between western and eastern Cherokee and within each, many of which predate the forced marched of the Cherokee from North Carolina to Oklahoma in 1838.

Northern Iroquoian consists of several subgroups. The first to break away from the main branch developed into Tuscarora, Nottoway, and Meherrin. The Tuscarora were first encountered in eastern North Carolina, but early in the 18th century most moved north to rejoin other Northern Iroquoians. Their descendants now live primarily in two communities, one at Six Nations on the Grand River in Ontario, the other near Niagara Falls in New York State. Nottoway and Meherrin were once spoken near the Virginia and North Carolina coasts. All that remains of Nottoway are two wordlists recorded during the early 19th century. The only record of Meherrin is two town names, sufficient to identify the language as Iroquoian. The Meherrin merged with those Tuscarora who did not migrate north in the 18th century.

The next group to separate from Northern Iroquoian became the Huron. They comprised a confederacy of four nations totaling around 20 000 people when they encountered Champlain in 1615 in present southern Ontario. In 1649, they were decimated by the Five Nations Iroquois. Some survivors formed a settlement at Lorette near modern Quebec City. Others joined the remnants of neighboring Iroquoian nations, the Tionontati (Petun), Erie, and Neutral, and migrated west toward Detroit and ultimately into Oklahoma. This group became known as the Wyandot. Both Huron proper at Lorette and Wyandot in Oklahoma were last spoken in the 20th century.

The remaining Northern Iroquoians separated into several subgroups, whose territories extended across present New York State. These were the Five Nations (with Tuscarora, the Six Nations), members of the League of the Iroquois. To the west were the Seneca and Cayuga. In the center were the Onondaga, near

Clift R (1999). 'Irony in conversation.' *Language in Society 28(4),* 523–553.
Gibbs R W, O'Brien J E & Doolittle S (1995). 'Inferring meanings that are not intended: speakers' intentions and irony comprehension.' *Discourse Processes 20,* 187–203.
Giora R (1995). 'On irony and negation.' *Discourse Processes 19,* 239–264.
Giora R & O'Fein O (1999a). 'Irony: context and salience.' *Metaphor and Symbol 14(4),* 241–257.
Giora R & O'Fein O (1999b). 'Irony comprehension: the graded salience hypothesis.' *International Journal of Humor Research 12(4),* 425–436.
Goffman E (1979). 'Footing.' *Semiotica 25,* 1–29.
Goldstein J H & McGhee P (1972). *The psychology of humor.* New York: Academic Press.
Haiman J (1990). 'Sarcasm as theater.' *Cognitive Linguistics 1(2),* 181–205.
Kaufer D S (1981). 'Understanding ironic communication.' *Journal of Pragmatics 5,* 495–510.
Leech G & Short M H (1981). *Style in fiction.* London: Longman.
Levinson S C (1983). *Pragmatics.* Cambridge: Cambridge University Press.
Martin R (1992). 'Irony and universe of belief.' *Lingua 87,* 77–90.
Muecke D C (1970). *Irony: the critical idiom.* London: Methuen.
Muecke D C (1978). 'Irony markers.' *Poetics 7,* 363–375.
Myers R A (1977). 'Towards a definition of irony.' In Fasold R W & Shuy R W (eds.) *Studies in language variation.* Washington, DC: Georgetown University Press. 171–183.
Myers R A (1981). 'The function of irony in discourse.' *Text 1(4),* 407–423.
Nash W (1985). *The language of humour: style and technique in comic discourse.* London: Longman.
Nilsen A & Nilsen D (2000). *Encyclopedia of humor and comedy.* Westport, CT: Greenwood Press.
Norrick N (1993). *Conversational joking.* Bloomington: Indiana University Press.
Raskin V (1985). *The semantic mechanisms of humor.* Dordrecht: D. Reidel.
Sperber D & Wilson D (1981). 'Irony and the use–mention distinction.' In Cole P (ed.) *Radical pragmatics.* New York: Academic Press. 295–318.
Triezenberg K (2004). 'Humor enhancers in the study of humorous literature.' *International Journal of Humor Research 17(4),* 411–418.
Wilson D & Sperber D (1992). 'On verbal irony.' *Lingua 87,* 53–76.

Iroquoian Languages

M Mithun, University of California, Santa Barbara, CA, USA

The Iroquoian Family

The Iroquoian languages are indigenous to southeastern and northeastern North America. The family consists of two major branches: Southern Iroquoian and Northern Iroquoian.

Southern Iroquoian is represented by just one language, Cherokee, spoken today primarily in Oklahoma and western North Carolina. There are clear dialect differences between western and eastern Cherokee and within each, many of which predate the forced marched of the Cherokee from North Carolina to Oklahoma in 1838.

Northern Iroquoian consists of several subgroups. The first to break away from the main branch developed into Tuscarora, Nottoway, and Meherrin. The Tuscarora were first encountered in eastern North Carolina, but early in the 18th century most moved north to rejoin other Northern Iroquoians. Their descendants now live primarily in two communities, one at Six Nations on the Grand River in Ontario, the other near Niagara Falls in New York State. Nottoway and Meherrin were once spoken near the Virginia and North Carolina coasts. All that remains of Nottoway are two wordlists recorded during the early 19th century. The only record of Meherrin is two town names, sufficient to identify the language as Iroquoian. The Meherrin merged with those Tuscarora who did not migrate north in the 18th century.

The next group to separate from Northern Iroquoian became the Huron. They comprised a confederacy of four nations totaling around 20 000 people when they encountered Champlain in 1615 in present southern Ontario. In 1649, they were decimated by the Five Nations Iroquois. Some survivors formed a settlement at Lorette near modern Quebec City. Others joined the remnants of neighboring Iroquoian nations, the Tionontati (Petun), Erie, and Neutral, and migrated west toward Detroit and ultimately into Oklahoma. This group became known as the Wyandot. Both Huron proper at Lorette and Wyandot in Oklahoma were last spoken in the 20th century.

The remaining Northern Iroquoians separated into several subgroups, whose territories extended across present New York State. These were the Five Nations (with Tuscarora, the Six Nations), members of the League of the Iroquois. To the west were the Seneca and Cayuga. In the center were the Onondaga, near

modern Syracuse. To the east were the Oneida and the Mohawk. The languages are now mutually unintelligible, though they share many structural features. Seneca is now spoken in three communities in western New York: Cattaraugus, Allegany, and Tonawanda. Most Cayuga left New York State after the Revolutionary War, some going to Six Nations in Ontario, and others to Oklahoma. The language is now spoken primarily at Six Nations. It was last spoken in Oklahoma late in the 20th century. Onondaga is spoken at Six Nations and at Onondaga south of Syracuse. Oneida is spoken near London, Ontario, and Green Bay, Wisconsin. There are six Mohawk areas: Kahnawà:ke and Kanehsatà:ke in Quebec; Ahkwesáhsne, which straddles Quebec, Ontario, and New York State; and Ohswé:ken (Six Nations), Thaientané:ken, and Wáhta' in Ontario. Most speakers now live in the first three.

Several other Iroquois languages are known to have existed as well. In 1534, Jacques Cartier encountered people along the St Lawrence River around present Quebec City. Vocabulary in his ship's logs and appended wordlists indicate that these people, now known as the Laurentian, spoke several Northern Iroquoian languages, at least one of which was not ancestral to any of the modern languages. When Champlain returned to the area in 1603, these people had disappeared. They are, however, the source of the name Canada. Another group, the Susquehannock or Andaste, were encountered by Captain John Smith about 1615 in the lower Susquehanna Valley in Pennsylvania. Their language, known through a wordlist recorded by the Swedish missionary Johan Campanius in 1696, was last spoken in the mid-18th century.

Phonology

The consonant inventories of the languages generally consist of one series of obstruents (reflexes of *t, *k, *tʃ, *kʷ, *s), one of resonants (*n, *r, *w, *j), and one of laryngeals (*h, *ʔ). Voicing is not distinctive. Noteworthy is the lack of labials. Most of the languages have four oral vowels (*i, *e, *a, *o) and two nasal vowels (*ę, *ǫ). In Proto-Northern-Iroquoian, stress was penultimate, and open, stressed syllables were lengthened.

There have been several innovations of interest. One is stress in Cayuga. Cayuga stress placement depends on syllable count from both edges of the word. Primary stress falls on the penultimate syllable, providing that it is even-numbered counting from the beginning of the word: *satkáhthoh* 'Look!' If the penult is odd-numbered, it is stressed only if it is open: *senhotǫ́:koh* 'Open the door!' If the penult is odd-numbered and closed, stress is antepenultimate: *sasaky'atáwihsih* 'Take your shirt off!' Epenthetic vowels, added after the establishment of basic penultimate stress in Proto-Northern Iroquoian, are not counted, like the *-a-* stem joiner in *sakyá'tawi't* 'Put your shirt on!' Stress on the second of two adjacent vowels moves to the first: *sasana̲háowe:k* 'Put your hat back on!' If conditions are not met for penultimate stress, and there is no antepenultimate syllable, the word carries no stress: *sa̲hsyǫ'* 'you returned.'

Cayuga shows another interesting innovation. In odd-numbered syllables closed by a laryngeal, the laryngeal feature spreads leftward over the entire syllable. If the laryngeal is h, the full syllable is devoiced. (Devoicing is shown orthographically by underlining.) If the laryngeal is glottal stop, the full syllable carries creaky voice, indicated here by a wavy underline: *w̲a̲hsí'ta̰'keh* 'on its foot.'

Several of the languages have developed distinctive tone under the effect of laryngeals. In Mohawk, stressed syllables generally show high pitch (if short) or rising pitch (if long). Stressed syllables closed by a laryngeal (glottal stop, or h followed by a resonant), however, developed a special pitch contour: a rise followed by a deep fall. The triggering laryngeal subsequently disappeared before a consonant. There are now contrasts, such as *oká:ra'* 'story' (with rising tone) and *okà:ra'* 'eye' (**okáhra'*). (Mohawk examples are given here in the community orthography. Nasalized vowels ʌ̨ and ų are written en and on respectively, glottal stop with an apostrophe ', vowel length with a colon :, stress accompanied by high or rising tone with an acute accent, and stress accompanied by falling tone with a grave accent. The palatal glide j is represented by i. Other symbols have approximate IPA values.) Western Cherokee underwent more complex changes resulting in tone. Feeling (1975) distinguishes three level tones (2, 3, 4), a rising tone (23), and two falling tones (32, 21, written 1), as in *a¹hyv²³ki³di²ýa* 'he's capturing him.'

Morphology and Syntax

All of the languages are polysynthetic. Verbs in particular can be composed of large numbers of meaningful parts (morphemes), like Mohawk *en-hske-rhar-átst-en-'* 'you will promise me' (FUT-2. sing/1.sing-expect-CAUS-BEN.APPL-PER). All verbs contain at least three parts: a pronominal prefix, a verb root, and an aspect suffix. The pronominal prefix refers to the core arguments of the clause: *-hske-*'you/me' in the verb above. The prefix appears in the verb whether or not there are coreferential independent nominals in the clause as well, but it is fully referential in its own right.

There is little evidence of Subject or Object categories. The pronominal prefixes show grammatical Agent/Patient patterning. The semantic basis underlying the system is still clear. Agent prefixes typically represent participants who are actively in control and instigating events or states, as in Mohawk ***ra**-tahónhsatats* '**he** is listening,' ***ra**-ná:ie* '**he** is conceited,' or ***rá**-hsere's* '**he** is chasing it.' Patient prefixes typically represent participants who are affected by a situation but not in control: ***ro**-thón:te'* '**he** hears it,' ***ró**-ta's* '**he** is sleeping,' or ***ró**-hsere's* 'it is chasing **him.**' The system is now fully routinized, however. Speakers do not judge degrees of control as they speak; they simply learn which set of prefixes to use with each verb.

Agent prefixes are used in both events and states (*jump, be conceited*), and Patient prefixes are used in both events (*holler*) and states (*be poor*). The system is thus basically Agent/Patient rather than Active/Stative. Superimposed on it, however, is one element of Active/Stative patterning. Some verbs occur only in the Stative: *rahseró:hen'* 'he is quick-tempered,' *ró:ten* 'he is poor.' Others occur in all three aspects: Habitual *ratè:kwahs* 'he escapes,' Perfective *wahatè:ko'* 'he escaped,' Stative *rotè:kwen* 'he has escaped.' With verbs like *escape*, which describe a change of state, the Stative aspect form typically has a Perfect meaning: *he has escaped*. All Perfect Stative verbs occur with Patient prefixes, whether their Habitual and Perfective forms appear with Agents or Patients.

The Northern Iroquoian languages show extensive, productive noun incorporation. The Southern language Cherokee shows traces of incorporation embedded in the lexicon, indicating that incorporation was productive in Proto-Iroquoian. Incorporation is a process whereby a noun stem is compounded with a verb stem to form a larger verb stem: *-itsker-onti-* 'saliva-throw = to spit.' There is no explicit specification of the semantic role of the incorporated noun; it simply indicates the involvement of a type of entity, often as a semantic patient, but sometimes as an instrument, location, source, or goal. Examples can be seen in Mohawk *wa-ho-n-**itsker**-ón:ti-'* FACTUAL-MASC.sing.AGT-MIDDLE-**saliva**-throw-PER 'he **saliva**-threw = he spit' and *ka-**hseriie't**-áner-en'* NEUTER-**cord**-tie-STATIVE 'it is **cord**-tied.' Incorporation is used pervasively to create new vocabulary and also to manipulate the flow of information through discourse. Important new participants are typically introduced with independent nominals, but those that are already part of the scene or of peripheral importance may be carried along as incorporated nouns.

Lexical categories are clearly distinguished by their internal morphological structure. Particles, by definition, are morphologically unanalyzable (*tóka'* 'maybe,' *oh* 'what,') though they may be compounded (*nek tsi* 'the only' for 'because'). Nouns contain a prefix specifying the gender of the referent or its possessor, a noun stem, and a noun suffix: *ka-nà:tsi-a'* NEUTER-kettle-NOUN.SUFFIX 'kettle,' *ake-nà:tsi-a'* 1.sing.ALIENABLE.POSS-kettle-NOUN.SUFFIX 'my kettle.' Alienable and inalienable possession are distinguished. Verbs follow an entirely different pattern, and can be quite complex morphologically. In addition to the obligatory pronominal prefix, verb root, and aspect suffix, they may contain various prepronominal prefixes. In the Northern languages, these are the Partitive, Contrastive, Coincident, Negative, Translocative, Factual, Duplicative, Future, Optative, Cislocative, and Repetitive. They may contain a Middle, Reflexive, or Reciprocal prefix. As seen above, they may contain an incorporated noun stem. Following the verb root, there may be one or more derivational suffixes, such as an Inchoative, Reversive, Causative, Instrumental Applicative, Benefactive Applicative, Distributive, or Purposive. After the aspect suffix, there may be a postaspectual suffix: a Past, Continuative, or Progressive. Such structure can be seen in the Mohawk *a-khe-'nikonhr-áks-a't-e'* OPT-1.sing/INDEF-mind-be.bad-CAUS-PER 'I would insult someone,' or *ia'-te-iako-hah-a-hiia'k-on-hátie'* TRANSLOCATIVE-DUPLICATIVE-FEM.PAT-road-STEM. JOINER-cross-STATIVE-PROG 'she was crossing the street.'

The morphological composition of particles, nouns, and verbs is thus entirely distinct. Noun stems never appear in the verb stem position of verbs, and verb stems never appear in the noun stem position of nouns. The match between internal morphological form and external syntactic function, however, is not isomorphic. Some morphological particles function syntactically and semantically as nominals, such as *è:rhar* 'dog.' Morphological verbs can function syntactically as predicates, as nominals (without further derivation), or as full clauses, as below.

Nahón:ne'	tehniiáhse
n-a-honn-e-'	te-hni-iahse
PART-OPT-MASC.PL.AGT-go-PER	DUPLICATIVE-MASC.DU.AGT-be.together.STAT
'they would go there'	'they two (males) are together'
niristì:sere's	nahshakotihahónnien'.
ni-rist-i'ser-e'-s	ne = a-hshakoti-hah-onni-en-'
MASC.DU.AGT-steel-drag-STAT-DIST	the = OPT-MASC.PL/3.DPL-road-make-BEN.APPL-PER
they (males) steel drag around	the they would road make for them

'They would go serve as guides for two surveyors.'

Finally, there is no basic, syntactically defined word order. In part because of the richness of the verbal morphology, the proportion of verbs to nouns and of predicates to nominals is much higher in Iroquoian languages than in many languages of Europe and Asia. There are few oblique or adjunct nominals. When clauses do contain multiple constituents, all orders are not only possible, they can also all be seen to be pragmatically motivated by the discourse at hand.

See also: Oneida; United States of America: Language Situation.

Bibliography

Chafe W (1967). Seneca morphology and dictionary. *Smithsonian Contributions to Anthropology 4*. Washington, D.C.

Feeling D (1975). *Cherokee–English dictionary.* Tahlequah, OK: Cherokee Nation of Oklahoma.

Foster M (1974). *National Museum of Man Mercury Series 20: From the earth to beyond the sky: an ethnographic approach to four longhouse Iroquois speech events.* (Cayuga). Ottawa: Canadian Ethnology Service.

Lounsbury F (1953). *Oneida verb morphology.* New Haven: Yale University Press.

Michelson K & Doxtator M (2002). *Oneida–English and English–Oneida dictionary.* Toronto: University of Toronto Press.

Mithun M (1979). 'Iroquoian.' In Campbell L & Mithun M (eds.) *The languages of Native America.* Austin, TX: University of Texas. 133–212.

Mithun M (1999). 'Iroquoian family.' *The languages ofNative North America.* Cambridge: Cambridge University Press. 418–430.

Mithun M & Henry R (1982). *A Cayuga grammar.* Brantford, ON: Woodland Indian Cultural and Educational Centre.

Rudes B (1999). *Tuscarora–English English–Tuscarora dictionary.* Toronto: University of Toronto Press.

Woodbury H (2003). *Onondaga–English English–Onondaga dictionary.* Toronto: University of Toronto Press.

Isixhosa *See:* Xhosa.

Islam and Arabic

M Abdel Haleem, School of Oriental and African Studies, London, UK

Arabic today is the main language of nearly 300 million people in the Middle East and North Africa. Since 1974 it has been recognized as one of the official languages of the United Nations. Like Hebrew, Aramaic, and Syriac, Arabic belongs to the Semitic group of languages. In the course of history it has been the language of a religion, a transnational state, and a widespread centuries-long civilization that has produced a rich heritage of knowledge, wisdom, and culture. It is used for religious purposes by 1.3 billion Muslims worldwide, and it is the language in which Islamic knowledge at its highest level has always been conducted. This status was achieved as a result of its connection to Islam through its sacred scripture, the Qur'an. The effect of the Qur'an on the Arabic language was, to an extent unknown for any other language of the world, dramatic, profound, and lasting.

The Impact of the Qur'an

Before the spread of Islam, Arabic was confined to the Arabian Peninsula, where it was spoken and used as a literary language by the native Arab population. Very little written material survives from the early centuries of the Christian era, the earliest being the Namara inscription in the Syrian desert, dated 328 C.E.; the output of the Arabs, particularly in poetry, was preserved as an oral tradition. In fact the Qur'an, when it was compiled in one volume, about 14 months after the death of the Prophet Muhammad (632) under his successor Abu Bakr, was the first book in Arabic. Thereafter, however, in line with the first revelation of the Qur'an in 610 C.E.:

> Read! In the name of your Lord who created: …. Read! Your Lord is the Most Bountiful One who ***taught by***

Finally, there is no basic, syntactically defined word order. In part because of the richness of the verbal morphology, the proportion of verbs to nouns and of predicates to nominals is much higher in Iroquoian languages than in many languages of Europe and Asia. There are few oblique or adjunct nominals. When clauses do contain multiple constituents, all orders are not only possible, they can also all be seen to be pragmatically motivated by the discourse at hand.

See also: Oneida; United States of America: Language Situation.

Bibliography

Chafe W (1967). Seneca morphology and dictionary. *Smithsonian Contributions to Anthropology 4*. Washington, D.C.

Feeling D (1975). *Cherokee–English dictionary.* Tahlequah, OK: Cherokee Nation of Oklahoma.

Foster M (1974). *National Museum of Man Mercury Series 20: From the earth to beyond the sky: an ethnographic approach to four longhouse Iroquois speech events.* (Cayuga). Ottawa: Canadian Ethnology Service.

Lounsbury F (1953). *Oneida verb morphology.* New Haven: Yale University Press.

Michelson K & Doxtator M (2002). *Oneida–English and English–Oneida dictionary.* Toronto: University of Toronto Press.

Mithun M (1979). 'Iroquoian.' In Campbell L & Mithun M (eds.) *The languages of Native America.* Austin, TX: University of Texas. 133–212.

Mithun M (1999). 'Iroquoian family.' *The languages ofNative North America.* Cambridge: Cambridge University Press. 418–430.

Mithun M & Henry R (1982). *A Cayuga grammar.* Brantford, ON: Woodland Indian Cultural and Educational Centre.

Rudes B (1999). *Tuscarora–English English–Tuscarora dictionary.* Toronto: University of Toronto Press.

Woodbury H (2003). *Onondaga–English English–Onondaga dictionary.* Toronto: University of Toronto Press.

Isixhosa *See:* Xhosa.

Islam and Arabic

M Abdel Haleem, School of Oriental and African Studies, London, UK

Arabic today is the main language of nearly 300 million people in the Middle East and North Africa. Since 1974 it has been recognized as one of the official languages of the United Nations. Like Hebrew, Aramaic, and Syriac, Arabic belongs to the Semitic group of languages. In the course of history it has been the language of a religion, a transnational state, and a widespread centuries-long civilization that has produced a rich heritage of knowledge, wisdom, and culture. It is used for religious purposes by 1.3 billion Muslims worldwide, and it is the language in which Islamic knowledge at its highest level has always been conducted. This status was achieved as a result of its connection to Islam through its sacred scripture, the Qur'an. The effect of the Qur'an on the Arabic language was, to an extent unknown for any other language of the world, dramatic, profound, and lasting.

The Impact of the Qur'an

Before the spread of Islam, Arabic was confined to the Arabian Peninsula, where it was spoken and used as a literary language by the native Arab population. Very little written material survives from the early centuries of the Christian era, the earliest being the Namara inscription in the Syrian desert, dated 328 C.E.; the output of the Arabs, particularly in poetry, was preserved as an oral tradition. In fact the Qur'an, when it was compiled in one volume, about 14 months after the death of the Prophet Muhammad (632) under his successor Abu Bakr, was the first book in Arabic. Thereafter, however, in line with the first revelation of the Qur'an in 610 C.E.:

> Read! In the name of your Lord who created: …. Read! Your Lord is the Most Bountiful One who ***taught by***

> *[means of] the pen,* who taught man what he did not know. (Q. 96:1–5)

the Arabs began to use the pen in earnest.

Before Islam, the Arabic script, derived from a Nabatean version of Aramaic, was rudimentary, containing only the consonants, and even among these there was no distinction between groups of letters of the same shape such as those representing *b, t,* and *th* or *j, h,* and *kh*. The script was developed in the Islamic era in order to achieve accurate pronunciation of the words of the Qur'an. Vowel signs and diacritical dots were introduced mainly by Al-Khalil ibn Ahmad Al-Farahidi (d. 786 C.E.), as commanded by the Ummayad governor Al-Hajjaj ibn Yusuf. This development continued further into the refinement of Arabic calligraphy as the noblest art of Islam, all to write and celebrate the word of God.

Within the space of 100 years after the death of the Prophet Muhammad in 632 C.E., Arabic had spread outside the Arabian Peninsula, both with the Arab armies and with emigration, into the Levant, the Fertile Crescent, Egypt, North Africa, and Spain, and in the East to Persia and beyond. It was to supersede, to a greater or lesser extent, such local languages as Aramaic, Greek, Coptic, Berber, and Persian, though naturally this process took some time. Arabic was first used for religious, administrative, political, and high cultural purposes and later spread to the population at large in the Middle East and North Africa, to become permanently the language of the majority. Persians wrote in Arabic for the first three centuries after Islam spread to Persia, and although Persian gradually regained precedence, Arabic left behind a great contribution of vocabulary and phraseology, particularly in religious and cultural fields. The Arabic script was also adopted for many Islamic languages, including Persian and later Urdu, Punjabi, Turkish, Hausa, Swahili, Malay, and Indonesian. Later Turkey changed, in 1928 and under European influence, to using the Roman script, as did many other countries, including Malaysia and Indonesia. Even in these countries, however, Arabic with its script continued to be used for religious studies.

Arabic Linguistic Scholarship

The Arabs did not study the grammar of their language before the Qur'an: it was in order to protect and understand the text of the Qur'an that Arabic grammar began to be codified. This started a flourishing tradition of Arabic linguistic scholarship. The important linguistic studies produced therein still command the admiration of modern linguists. Arabic phonetics, which started in order to describe exactly the proper pronunciation and articulation of the Qur'an, developed into a science of *tajwid,* good recitation of the Qur'an, an art that continues to develop. Similarly, the study of Arabic syntax flourished for the service of the Qur'an.

Especially significant in this field are works of the outstanding scholar, Abdul Qahir ibn Abdul Rahman al Jurjānī (d. 1078 C.E.), who dedicated his *Dalā'il al-I'jāz* and his *Asrār al-balāgha* to investigating the style and rhetoric of the Qur'an, in an attempt to identify the reasons for the inimitability (*i'jāz*) of its language. To give a brief example, two key concepts were discussed by al-Jurjani and others: context and internal relationships. The importance of context (*maqām*) was recognized and formulated for the study of the text of the Qur'an by Muslim linguists. In this respect, their work anticipated by many centuries modern linguistic thinking about the crucial importance of context in understanding discourse. Internal relationships were encapsulated in the dictum: *al-Qur'an yufassir ba'duhu ba'da* (different parts of the Qur'an explain one another), which, given the structure of Qur'anic material, was argued to provide the most correct method of understanding the Qur'an.

The text of the Qur'an was crucial for the development of the Arabic science of rhetorics, but Arab writers also drew on the text of the Qur'an in their study of the principles of literary criticism. This can be seen in the celebrated work of the famous critic Ibn al-Athir in his book *Al-mathal al-sā'ir fī adab al-kātib wa sha'ir* ('On the criticism of prose and poetry'). Lexicography was another discipline that was started to serve the Qur'an and flourished to a great extent in the Arabic tradition, the first, a landmark dictionary, being *Kitab al-'ayn* by Al-Khalil Ibn Ahmad (died 786 C.E.). As stated by Suyuti (d. 1505 C.E.), a writer on many subjects including Qur'anic studies and linguistics, the Qur'an was in fact the source for all branches of Arab knowledge.

The status of linguistic studies in the Arabic tradition was elevated because of its connection with the Qur'an. In our first grammar lessons, at the primary school of Al-Azhar in Cairo (the oldest functioning institution of Islamic learning in the world, opened in 972 C.E.), we were taught *Al-Mabadi' al-'ashara,* The Ten Introductory Aspects, with which the textbook began; we then learned that these ten aspects, as handed down by the ancient tradition, were points to be considered when embarking on any branch of knowledge. They included the definition of that particular branch, its name, its issues, its founder, its relation to other subjects, its status among other subjects, and how the Shari'ah views the learning and application of that subject. The status of grammar, we were told, was very high because it ensures correct

understanding and reading of the Arabic text of the Qur'an. As such its study was a communal obligation on the Muslim community. Thus learning grammar and linguistics, in addition to any educational or secular purpose, was a religious duty in Al-Azhar and similar traditional institutions of learning in the Muslim world.

Expansion of Arabic Knowledge

In the early centuries of Islam, the Muslims, inspired by the Qur'anic injunction, 'Say, "Lord increase me in knowledge"' (20:114), considered it their duty to translate (into Arabic), absorb, and develop the valued knowledge of ancient civilizations, such as Greece, Rome, Persia, India, and China. This provided another channel by which the local literary output in Arabic grew under Islam. In this case it drew in external sources to form another part of the vast corpus of learning in Arabic. Some of this Arabic scholarship was then, from the 12th century C.E., translated into Latin and acted as a contributing factor to the European Renaissance. The great tradition of learning in Arabic was developed not only by Arabs but also by Muslims from all over the world in their service of Islam. Chief among these were scholars from Persia and Central Asia, writing in Arabic. Worthy of special mention are scholars of Qur'anic exegesis (e.g., Al-Rāzī, d. 1209), the compilation and criticism of Prophetic sayings (hadīth) (Bukhārī, d. 870; Abū Dāwūd, d. 888; Tirmidhī, d. 892; and Al-Nasā'ī, d. 916), Arabic linguistics (e.g., Sibawayhi, d. 793 and Al-Jurjanī, d. 1078), and philosophy (Ibn Sīnā, d. 1037 and Al-Farābī, d. 950).

Arabic Studies in the West

We may mention here another vigorous tradition of Arabic linguistic scholarship, this time in Western universities (especially since the 17th century). This tradition was, from the beginning, driven above all by its relevance to Islam, and it produced monumental works in the study of Arabic and Islam in the leading centers of learning and in various European languages such as Dutch, English, French, German, Italian, Russian, and Spanish.

Memorizing the Qur'an, and Its Use in Worship

It is the Arabic Qur'an that is considered to be the word of God, not any translation. This makes Muslims throughout the world eager to learn some of it in Arabic. Both Arabs and non-Arabs, including women, are known to memorize the whole Qur'an or a large amount of it. There are international Qur'an recitation competitions held in Malaysia, Indonesia, and Iran, to name but a few, sometimes state funded, where the Qur'an is recited from memory. There is one particular passage, the first Sura, *Al-Fātiha*, which all Muslims learn, as it is an obligatory part of the five daily prayers in Islam, and normally, in addition to this, they learn other short passages. Muslims all over the world perform their daily prayers in Arabic, in imitation of the Prophet's example, and in keeping with the unbroken tradition, even though, according to at least one school of law, the Hanafi, if they do not know Arabic, they can perform the prayer in their own language. If we assume a group of ten Muslims come together from differing linguistic backgrounds, with no common language, they can all pray together, as it is the same Arabic prayer for all of them. Any of them is, moreover, eligible to lead the prayer.

As the Qur'an and Hadith are in Arabic and are the fundamental sources of Islamic law and teaching, the main classical works of the different schools of law, which are followed to this day, are all written in Arabic. This means that advanced Islamic studies in non–Arab Muslim countries are conducted in Arabic by the followers of all the various schools, Shāfi'ī in South East Asia, for instance, Hanafī in South Asia and Turkey, Ja'afarī in Iran, and Mālikī in West Africa. Thus Islam has kept Arabic alive in the religious scholarship of non-Arab Muslims. Even within the central area of the Islamic world, the Arab Middle Eastern and North African countries, it is Islam that has kept the Arabic language alive and given it this unrivaled position. The literary language, common to all Arabs, is used for writing, education, and formal speech. The connection of Arabic to the Qur'an has secured this position for the Arabic language up to today. Around the beginning of the 20th century in Egypt, for instance, there were calls for the adoption of English or of the Roman script for writing, or even the local Arabic dialect of Egypt. These were fiercely rejected because it would have meant cutting off the following generations from the language of the Qur'an and Hadith, Islamic Law, and the Arabic literary heritage.

Arabic in Western Cities

In recent times, Muslims have settled in every continent, including Europe and America. Arabic schools are set up for their children, also mosque schools, where the Arabic Qur'an is taught. It is remarkable that Muslims from non-Arab countries such as South Asia, Turkey, and Africa insist on having a Friday sermon in Arabic from the pulpits in Europe and

America, even if they have additional sermons in their own languages normally preceding the prayer time. Despite the growing body of Muslim writing on Islam in European languages, Arabic, the language of the Qur'an, still has a privileged position that cannot be replaced by any other language.

See also: Al-Jurjani, Abd Al-Qahir (10th Century A.D.); Al-Khalil (8th Century A.D.); Al-Zajjaji, Abu Al-Qasim (10th Century A.D.); Arabic Linguistic Tradition; Arabic; Islam in Africa; Islam in East Asia; Islam in Southeast Asia; Islam in the Former Soviet Union; Muslim/Islamic Language Education in Pakistan and India; Qur'an; Semitic Languages; Sibawayhi (8th Century A.D.).

Bibliography

Abdel Haleem M A S (1991). 'Early Islamic theological and juristic terminology: *Kitāb al-Hudūd fī al-Usūl* by Ibn Fūrak.' *Bulletin of the School of Oriental and African Studies 54,* 1.

Abdel Haleem M A S (2001). *Understanding the Qur'an: themes and style.* London: I. B. Tauris.

Abdel Haleem M A S (2004). *The Qur'an: a new translation* (with introduction). Oxford: Oxford University Press.

Abu Deeb K (1979). *Al-Jurjani's theory of poetic imagery.* Warminster: Aris.

Bakalla M H (1982). *Arabic linguistics: an introduction and bibliography.* London: Mansell.

Faruqi I R & Faruqi L L (1986). *The cultural atlas of Islam.* New York: Macmillan.

Gibb H A R *et al.* (eds.) (1960). Entries on '*arabiyya*', 'Islam', and '*khatt*'. In Gibb H A R *et al.* (eds.) *Encyclopaedia of Islam* (2nd edn.). Leiden: E. J. Brill.

Jones R (ed.) (1978). *Indonesian etymological project III: Arabic loan words in Indonesian.* London: School of Oriental and African Studies.

Kratz E U (1999). 'Islam and Indonesia.' In Clark P B (ed.) *The world's religions: Islam.* London: Routledge Reference.

Safadi Y H (1978). *Islamic calligraphy.* London: Thames and Hudson.

Shackle C (2000). Entry on 'Urdu.' In Gibb H A R *et al.* (eds.) *Encyclopaedia of Islam, 9,* 2nd edn. Leiden: E. J. Brill.

Steiger A (1963). *Origin and spread of Oriental words in European languages.* New York: S. F. Vanni.

Islam in Africa

J Hunwick, Skokie, IL, USA

After the death of the Prophet Muhammad in 632 A.D., Arab Muslim forces overtook most elements of the Byzantine-ruled areas of the Middle East, and then in 639 moved into the African continent, taking over Egypt from the Byzantines, and then moving westward through North Africa. They reached what is now Tunisia and established their fortress town at Qairawan in 670, and by the end of the 7th century they had occupied land up to the coast of the Atlantic Ocean, though not taking over territory very far south of the Mediterranean Sea.

Importantly for them, the Berber populations of their occupied territories mostly converted to Islam. In this way, Islam passed on to Saharan nomads, particularly the Berber Sanhaja of the western Sahara. With a renewal and growth of their faith in the mid-11th century, the Sanhaja formed a movement, known as the Almoravids, that then occupied all of the region of modern Morocco, and moved into the area of Spain formally overtaken by Arab Muslims. As widespread Saharan nomads, their faith to Islam soon began to have influence on West African black populations. The Massufa clan of the Sanhaja migrated along the southern region of the Sahara in the late 11th century, some of them settling in Walata, Sane (near Gao), and the Takedda region; some Sanhaja took up an area that later became Timbuktu, and, indeed, it was Sanhaja scholars who later made Timbuktu a major center of Islamic learning.

Among early West Africans to adopt Islam were the Dyula (also called Wangara), a clan of the Mande ethnic group, the rest of which later became Muslims. The Dyula mainly engaged in trade and moved widely around West Africa, influencing other peoples to adopt Islam. Another important ethnic group to largely adopt Islam was the Fulani (or Fulbe), nomadic cattle owners, who moved from the edge of the southern Sahara down into the Inland Delta of the Nile river and then elements of this group moved to other areas, such as Futa Toro and Futa Jallon, and eventually moved eastward into northern Nigeria and on to Chad and Dar Fur. By the 9th century, North African Muslims (Arab/Berber) began to cross the Sahara for trade, including to obtain slaves, and no doubt they had influence of conversion on the rulers they contacted.

Among the early and most celebrated persons converted to Islam were a number of rulers in the Sahel zone: the ruler of Takrur, in the Senegal river area;

America, even if they have additional sermons in their own languages normally preceding the prayer time. Despite the growing body of Muslim writing on Islam in European languages, Arabic, the language of the Qur'an, still has a privileged position that cannot be replaced by any other language.

See also: Al-Jurjani, Abd Al-Qahir (10th Century A.D.); Al-Khalil (8th Century A.D.); Al-Zajjaji, Abu Al-Qasim (10th Century A.D.); Arabic Linguistic Tradition; Arabic; Islam in Africa; Islam in East Asia; Islam in Southeast Asia; Islam in the Former Soviet Union; Muslim/Islamic Language Education in Pakistan and India; Qur'an; Semitic Languages; Sibawayhi (8th Century A.D.).

Bibliography

Abdel Haleem M A S (1991). 'Early Islamic theological and juristic terminology: *Kitāb al-Hudūd fī al-Usūl* by Ibn Fūrak.' *Bulletin of the School of Oriental and African Studies 54,* 1.

Abdel Haleem M A S (2001). *Understanding the Qur'an: themes and style.* London: I. B. Tauris.

Abdel Haleem M A S (2004). *The Qur'an: a new translation* (with introduction). Oxford: Oxford University Press.

Abu Deeb K (1979). *Al-Jurjani's theory of poetic imagery.* Warminster: Aris.

Bakalla M H (1982). *Arabic linguistics: an introduction and bibliography.* London: Mansell.

Faruqi I R & Faruqi L L (1986). *The cultural atlas of Islam.* New York: Macmillan.

Gibb H A R *et al.* (eds.) (1960). Entries on '*arabiyya*', 'Islam', and '*khatt*'. In Gibb H A R *et al.* (eds.) *Encyclopaedia of Islam* (2nd edn.). Leiden: E. J. Brill.

Jones R (ed.) (1978). *Indonesian etymological project III: Arabic loan words in Indonesian.* London: School of Oriental and African Studies.

Kratz E U (1999). 'Islam and Indonesia.' In Clark P B (ed.) *The world's religions: Islam.* London: Routledge Reference.

Safadi Y H (1978). *Islamic calligraphy.* London: Thames and Hudson.

Shackle C (2000). Entry on 'Urdu.' In Gibb H A R *et al.* (eds.) *Encyclopaedia of Islam, 9,* 2nd edn. Leiden: E. J. Brill.

Steiger A (1963). *Origin and spread of Oriental words in European languages.* New York: S. F. Vanni.

Islam in Africa

J Hunwick, Skokie, IL, USA

After the death of the Prophet Muhammad in 632 A.D., Arab Muslim forces overtook most elements of the Byzantine-ruled areas of the Middle East, and then in 639 moved into the African continent, taking over Egypt from the Byzantines, and then moving westward through North Africa. They reached what is now Tunisia and established their fortress town at Qairawan in 670, and by the end of the 7th century they had occupied land up to the coast of the Atlantic Ocean, though not taking over territory very far south of the Mediterranean Sea.

Importantly for them, the Berber populations of their occupied territories mostly converted to Islam. In this way, Islam passed on to Saharan nomads, particularly the Berber Sanhaja of the western Sahara. With a renewal and growth of their faith in the mid-11th century, the Sanhaja formed a movement, known as the Almoravids, that then occupied all of the region of modern Morocco, and moved into the area of Spain formally overtaken by Arab Muslims. As widespread Saharan nomads, their faith to Islam soon began to have influence on West African black populations. The Massufa clan of the Sanhaja migrated along the southern region of the Sahara in the late 11th century, some of them settling in Walata, Sane (near Gao), and the Takedda region; some Sanhaja took up an area that later became Timbuktu, and, indeed, it was Sanhaja scholars who later made Timbuktu a major center of Islamic learning.

Among early West Africans to adopt Islam were the Dyula (also called Wangara), a clan of the Mande ethnic group, the rest of which later became Muslims. The Dyula mainly engaged in trade and moved widely around West Africa, influencing other peoples to adopt Islam. Another important ethnic group to largely adopt Islam was the Fulani (or Fulbe), nomadic cattle owners, who moved from the edge of the southern Sahara down into the Inland Delta of the Nile river and then elements of this group moved to other areas, such as Futa Toro and Futa Jallon, and eventually moved eastward into northern Nigeria and on to Chad and Dar Fur. By the 9th century, North African Muslims (Arab/Berber) began to cross the Sahara for trade, including to obtain slaves, and no doubt they had influence of conversion on the rulers they contacted.

Among the early and most celebrated persons converted to Islam were a number of rulers in the Sahel zone: the ruler of Takrur, in the Senegal river area;

the king of the Zuwa dynasty in Gao, and the ruler of Kanem, just north of Lake Chad. All of these conversions were in the 11th century. Later, rulers of major empires promoted Islam, making pilgrimages to Mecca, e.g., Mansa Musa of Mali in 1324 and Askiya Muhammad of the Songhay Empire in 1497. As a pious Muslim, Askiya gave support to the Islamic community of Timbuktu, a city of Muslims ever since its initiation in the early 12th century. Muslim scholars migrated there from North Africa and from Saharan oases, and when Mansa Musa returned from his pilgrimage, he stopped in Timbuktu and had the Great Mosque (Jingere Ber) built, a Muslim shrine soon afterward to be added to with the large Sankore mosque. The most celebrated scholarly clan to settle in Timbuktu was the Aqit family of the Sanhaja Massufa, coming from Masina in the south of the Inland Delta.

Many works on aspects of Islam – especially shariah – were written in Arabic in Timbuktu, especially in the 15th–17th centuries, and even down to the 20th century. Some Muslim scholars possessed large libraries of such manuscripts – many still in existence – and in 1973 a huge public library of manuscripts was founded: the Ahmad Baba Center (Centre de Documentation et de Recherches Historiques Ahmad Baba [CEDRAB]). Ahmad Baba (1556–1627), after whom it was named, was one of the most celebrated Timbuktu scholars, a member of the Aqit family, though he was exiled to Morocco in 1594, returning home in 1608. He wrote nearly 70 works, and became celebrated in Morocco as well as in West Africa.

A West African scholar who once quoted him was the Fulani Uthman ibn Muhammad Fudi (or Usman dan Fodio), who set up a jihad in the Sokoto region of northwestern Nigeria in 1804, creating an Islamic empire out of most of the Hausa states of what is now northern Nigeria. His ancestors had migrated to that region from Futa Toro in the 16th century, though Islam had other influences in Hausaland. Kanem became an early Islamic state, and in the 13th century took over what is now Bornu, and from there, to the west of Lake Chad, Islam then had influence farther west. There was also influence in one area of Hausaland from the Dyula. Some of them had moved south from Jenne to obtain the gold being mined in what is now southern Ghana, to be taken to Jenne and then Timbuktu for trans-Saharan trade. When Songhay took over the Inland Delta in the 1470s, the Dyula decided to take the gold to a different trans-Saharan trading location, and so some of them went off to Kano, others of them having already converted the ruler of that city and state, a trading zone with trans-Saharan interconnections.

Islam later spread farther south in West Africa. In the second half of the 17th century, some scholars moved from Timbuktu to the north of what is now the Republic of Ghana. In the same period, some Dyula also settled in that area, and a century or more later Hausa migrants moved into the region, so the area between the Black and White Volta rivers became a new region of Islam, and by the 19th century a source of Arabic-Islamic writings. Similarly, in what is now Nigeria, Islam's influence moved south, following the early 19th-century jihad that created the Sokoto Caliphate. The Nupe adopted Islam, and in 1824 the northern Yoruba town of Ilorin became part of the Caliphate, and soon produced Muslim scholars who wrote Arabic treatises. Some later moved south as far as Abeokuta and Lagos; notably Adam al-Iluri (1917–1992), who set up at Agege, near Lagos, an Arabic and Islamic training center in the mid-20th century, and published there many of his 70 Islamic writings and those of some of his colleagues.

Although West Africa became a major area of Islamic writings, Islam became the main religion of other areas of Africa. Many nomadic Arabs moved into Egypt south of Cairo, and by the 14th century began moving farther south, causing Islam to take over the Christian kingdom of Nubia in the north of the present Republic of the Sudan. In the following centuries, Arab Muslim nomads in Egypt took over much of what is now that republic, while others moved directly from Arabia across the Red Sea. In 1820–1821, the Sudan was taken over by Egypt, then part of the Ottoman Empire, and some 60 years later a north-Sudan Muslim scholar, Muhammad Ahmad ibn Abdullah, declared himself to be '[God-]guided' – the Mahdi, a position normally seen as the final reviver of God's will for humanity. Early in 1885, he took over the main city of Khartoum, after controlling all the rest of the Muslim-inhabited sectors of the country, and died a few months later, to be succeeded by a successor: the *khalifa*. Although in 1898 the Sudan was occupied by the British and the Egyptians, who fought and killed the Khalifa, it remained, except for the far south, a land of Islam, including Darfur, a Muslim state to the west, now on the border of the Chad Republic.

Bordering the Sudan to the east is the state of Ethiopia (plus Eritrea), which was originally a Christian land, but its lowlands became Islamic territories in the 14th century. Some Muslims from Mecca had fled to an area of northern Ethiopia in 615, and the Arabs later respected Ethiopians, even before some converted to Islam. Ethiopia then borders Somalia, now nearly 100% Muslim, with some of its tribes, although not ethnically fully Arab, claiming that

they migrated to the Horn of Africa from Arabia many centuries ago. Early in the 20th century, just as the British were beginning to occupy a northwestern area of the Horn, a Muslim man, Mohamed Abdille Hassan (1864–1920) referred to by some as a *mahdi*, brought together many other Somalis, attempting to get them to abandon the Qadiriyya Sufi order and adopt the Salihiyya order that he had adopted in a visit to Mecca, and also to reject colonialism. The Salihiyya Sufi order was a product of Sufi teachings promoted by Ahmad ibn Idris, who moved out of Morocco in 1789, along North Africa to Cairo and then Upper Egypt, and finally into Arabia – Mecca and then Yemen. His teachings produced a number of branches of his movement. Other than the Salihiyya, several parts of Africa received his disciples. Libya, for example, got Muhammad al-Sanusi (1787–1859), of Algerian origin, whose order, the Sanusiyya, took off in Jaghbub and then expanded into southern Libya, and later, under his successor, into much of Chad. Another disciple was Muhammad Uthman al-Mirghani (1793–1852), from a Meccan Sharifian family, who was sent by his teacher to Eritrea, where he was treated hostilely and returned to Mecca; then he went to Upper Egypt and later into the Sudan for 4 years, promoting *al-Mirghaniyya*; and when in Kordofan, a man called Isma'il al-Wali took over some of his teachings and set up his own order, the Isma'iliyya. Al-Mirghani's Sufi teachings then became known as al-Khatmiyya, and he promoted them elsewhere in the Muslim lands of the Sudan. Also, another branch of the Idrisiyya (the teachings of Ahmad ibn Idris) was taken down into East Africa, where the Swahili populations were already Muslims. The Swahili population and language had, in fact, come to exist through Arabian visitors and settlers from the early centuries of Islam.

Islam also became practiced much farther south – in what is now called South Africa – mainly by people of Asian origin, brought into the region by colonial authorities: by the Dutch in the 18th century, many as slaves, including some from West or East Africa; then in the late 19th century the British brought Indians as indentured laborers. Through their presence, some local black Africans converted to Islam. In Cape Town, several mosques were built in the center of the city, and recently the Koran has been translated into their Xhosa language.

As for Sufi orders – or *tariqas* as they are called in Arabic – two carried strong spiritual influence on western Africa. The first was the Qadiriyya, originating in Baghdad in the 12th century under the teachings of Abd al-Qadir al-Jilani, later adopted in North Africa and by Saharan tribes such as the Kunta, an Arab-Berber tribe containing intellectuals, originally in the Touat oasis, who moved south from the 17th century and split in two, with a major group just north of Timbuktu and governing Timbuktu in part of the first half of the 19th century. The Qadiriyya was also adopted in northern Nigeria (notably by Uthman ibn Muhammad Fudi) and in Senegal. Another *tariqa*, the Tijaniyya, overtook it in the 19th century, though not among the Kunta. Ahmad al-Tijani (1737–1815), the founder of that Sufi order, was born in an Algerian oasis Ayn Madi, and then moved to Fez in Morocco. His teachings – following study of other Sufi *tariqas* – were adopted by the Idaw Ali tribe of southern Mauritania, and were also promoted by al-hajj Umar ibn Sa'id of Futa Toro in Senegal, who later conquered much of present-day Mali. The Tijani *tariqa* became appreciated in Senegal, and led by the shaykh Ibrahim Niasse (1902–1975), it was adopted by many West Africans as he traveled into Mali, Ghana, and Nigeria, and subsequently on to Chad and the Sudan. In Senegal, another *tariqa* was created: the Mouride order, inspired by Ahmad Bamba (1855–1927), originally a Qadiri. In the second half of the 20th century, however, Sufi *tariqa*s – especially the Tijaniyya – were opposed in speech and writing by Africans who had spent time in Arabia and come under the conveying influence of the Wahhabiyya movement. Whereas Sufi orders put little emphasis on *shari'ah*, the Wahhabiyya created its own Islamic law views, based on the Hanafiyya, though all West African Muslims belonged to the Maliki *madh'hab*.

See also: African Linguistics: History; Berber; Chadic Languages; Hausa; Islam and Arabic; Language Education Policies in Africa; Mande Languages; Nilo-Saharan Languages; Qur'an; Somali; Songhay Languages.

Bibliography

Encyclopaedia of Islam (11 vols.) (1960–2003). Leiden: Brill.

Hunwick J & O'Fahey R S (eds.) (1994–2005). *Arabic literature of Africa* (vols. 1–4). Leiden: Brill

Holt P M & Daly M W (1979). *The history of the Sudan, from the coming of Islam to the present day.* Boulder: Westview Press.

Lewis I M (1965). *The modern history of Somaliland.* London: Weidenfeld & Nicholson.

O'Fahey R S (1990). *Enigmatic saint: Ahmad Ibn Idris and the Idrisi tradition.* London: Hurst & Company.

Islam in East Asia

J Wang, Shanghai Normal University, Shanghai, China

Generally, the Muslim population in China is scattered over the country but concentrated unevenly in certain areas; as a result, Islam in China shows a considerable diversity of detail, although the basic Islamic principles are preserved in the form of community. Chinese Muslims are a group of different ethnic groups, languages, customs, and cultures. Therefore, Islam in inland China is always in interaction with Chinese religions and Chinese cultures; however, Islam in the frontier regions of China is connected with its counterparts in Central Asia, South Asia, and South-East Asia.

The first bearers of Islam to China were not missionaries but merchants who came to China by the Silk Road on land and for the spice trade by sea. The official contacts of the Arab Caliphate and the Chinese Empire began in 651 as the Caliph Uthman sent an envoy to Xi'an, the capital under the Tang Dynasty (618–907). Another large wave of Muslim immigrants consisted of the warriors who were sent by the Caliph Jafar Mansur of the Abbasid Dynasty in 755 to help the Chinese empire quell a rebellion. After recapturing the capitals of China, these soldiers were given the right to stay in China as a gratuity for their service. Their settlements developed into Muslim communities.

The earliest period of Islam in China did not exercise much influence over the Chinese people. The Muslims lived in a special section of a town and kept their own customs and way of living, had their own administration system, built mosques, and appointed their own judge to be in charge of religious and legal affairs. In a riot in 878, the revolting Chinese peasants besieged Canton, in which about 120 000 Muslims, Christians, and Jews were killed and their communities looted. However, during most of this period, the relations between the Muslims and the Chinese were peaceful.

When the nomadic Mongols conquered China, they brought many Muslims in their troops from Central Asia and Persia, and so began the second influx of Muslims to China who settled. The Muslims were important allies and aides of the Mongols in ruling this empire and forced the large subject population to submit to the new authority. The sizeable Muslim groups became elites in Chinese society. Not a few Muslims were appointed by the Mongols to the posts of prime minister, minister, and governor, etc. Muslim scientists made contributions to Chinese culture on astronomy, mathematics, geometry, architecture, medicine, and the technologies of handicrafts and weaponry. The spread of Muslims all over the country took place during this Mongol Dynasty (1271–1368). When Marco Polo and Ibn Battuta traveled in China, both noticed that China had an impressive Muslim population, and many towns and cities had special Muslim resident sects, and there were even Sufi monasteries active in the communities.

In Chinese Turkistan, Islam follows another mode. In the 10th century, an Uighur king of the Qara Khan Dynasty under the influence of an older Sufi from Central Asia was converted to Islam and given a Muslim name, Sutuk Baghra Khan. About 200 000 Uighurs also followed him in conversion. The new converts waged Jihad to conquer the other neighboring Uighur kingdoms, which upheld Buddhism, Manichaeism, and Shamanism. However, the Islamization of this region was a long process, and after the turn of the 16th and 17th centuries the whole region called Xinjiang today became part of the Islamic world.

The privileges of Muslims continued during the Ming Dynasty (1368–1644), when the majority Han people resumed their sovereignty over China. The position of Muslims in China did not meet great change except that the first emperor of the dynasty prohibited the Muslims from exercising their 'foreign culture'. Islam in inland China began a rapid acculturation process. Many Muslims devoted themselves to learning Chinese literatures, and in this way they became officials of the civil and military services in the empire. Many Muslims lost their Central Asian identities in this period when they spoke Chinese language and wore local clothes, although some Persian and Arabic terms were kept for their Islamic services. However, the Turkic Muslims retained their Central Asian Islamic culture and ethnic identities due to their close contacts with that region.

In the Qing Dynasty (1644–1911), the relations between Muslims and non-Muslims and between Muslims and the imperial authorities deteriorated, which led to the rigorous policy of suppressing the Muslims. More than ten million Muslims lost their lives in the two major confrontations during the 18th and 19th centuries.

When the Republic (1911–1949) was founded, the Muslims were one of five nationalities in China with equal rights in politics. This change of status in economic, social, and cultural life strengthened contacts with other parts of the Islamic world and reawakened the national consciousness of the Muslim peoples. New-style Islamic schools were set up in many communities. A national network of Islamic associations

was established. But in the 1940s, discontented by the high-handed policy and worsening economic situation, the Turkic Muslims in Xinjiang rose up and founded the short-lived Eastern Turkistan Republic.

The Communist regime replaced the Republic in 1949. Under the extremist leftist influence, the 'Religious Reform' was launched in 1958. Many mosques were closed, and the clerics were put into labor camps. The hostels of the Sufi orders and saints' tombs were forced to be demolished. The religious rights of Muslims had never been threatened to such a degree as during the Cultural Revolution (1966–1976). All religious activities were stopped, and many believers were killed as counter-revolutionaries.

However, since the beginning of the 1980s, the new Chinese leaders have embarked upon a more pragmatic policy with regard to religious affairs. Islam has gained a dramatic revival. China has more than twenty million Muslims with more than 30 000 mosques, according to country's official statistics. Several sects such as Shi'a, Ikhwan Movement, Salafiyya, Jahriyya, Khufiyya, and Qadriyya, as well as Qadim, the traditional Sunni mainstream, are allowed free operation. Many Islamic schools have been established to train the young people in Islamic knowledge and tradition. Every year thousands of Muslims sponsor themselves to go to Mecca on pilgrimage. Many communities publish their own magazines and newspapers to promote Islamic doctrine. Occasional clashes occur between Muslims and non-Muslims or local authorities because of different interests. In Xinjiang, some Uighur Muslims inspired by the event of the newly independent Central Asian republics fight for separatism under the banner of Jihad. The Communist regime through closely monitored organizations such as the Association of Islam tries hard to make Islam accommodate to the Chinese economic-market-type socialism to keep social stability.

Islam affected the Korean peninsula at first through indirect contacts via China in the period of the 10th to the 13th centuries. The *Koryosa* chronicle informs us that Muslims came to Korea through commercial activities and were treated as honored guests by the Kingdom's government (Jeong, 1452). When the Mongols controlled Korea in 1270, some Central Asian Muslims served as administrative officials and settled down in the peninsula where they formed communities. However, in 1427 the government issued a decree prohibiting Islamic performance and customs, and so therefore mostly assimilated the immigrant Muslims. The emergence of new Muslim communities in Korea has taken place 500 years later, in the second half of the 20th century, as South Korea became a strong country economically and came to rely on oil from the Middle East, which led Muslim foreigners to to Korea for trade, study, and work. There is no exact figure for the Muslim population in Korea, but they are certainly not a big community, and mainly composed of Muslim immigrants and local converts.

Islam traveled to Japan about 1891–1922, and some local Japanese embraced this faith. This was followed by Tartar Muslim migrants from Central Asia in 1923 to 1952. In 1956, the Muslim missionaries from Pakistan and India started going to Japan, playing an important role in the spread of Islam. In the late 1950s, students from Muslim countries went to study at Japanese universities. In 1961, the Muslim Students Association was founded. In 1974, the Islamic Center was established in Japan. Several big Japanese cities including Tokyo have mosques supported by Islamic countries such as Saudi Arabia. The Japanese Muslim population is estimated at less than 50 000.

See also: China: Religions; Chinese; Islam in Southeast Asia; Japan: Language Situation; Jewish Languages; Korea, Republic of (South): Language Situation; Persian, Modern; Saudi Arabia: Language Situation.

Bibliography

Broomhall M (1910). *Islam in China – A neglected problem.* London: Morgan and Scott.

Chang B O (2003). 'Korea's foreign policy toward the Middle East.' *Annals of Japan Association for Middle East Studies 18(2).*

Gladney D C (1991). *Muslim Chinese: ethnic nationalism in the People's Republic of China.* Cambridge, MA: Harvard University Press.

Israeli R (1980). *Muslim in China: a study in cultural confrontation.* London: Curzon.

Jeong (1452). 'Islam in Korea.' In Ji (ed.). *Koryosa.* Seoul.

Korea Muslim Federation (1996). *Special issue on the 40th anniversary of Islam in Korea (1955–1995).* Seoul.

Leslie D D (1986). *Islam in traditional China.* Belconnen, Australia: Canberra College of Advanced Studies.

Lipman J (1997). *Familiar strangers.* Seattle: University of Washington Press.

Wang J (1996). *Concord and conflict: the Hui communities of Yunnan society in a historical perspective.* Stockholm: Almqvist and Wiksell International.

Wang J (2001). *Glossary of Chinese Islamic terms.* London: Curzon.

Islam in Southeast Asia

E U Kratz, SOAS, London, UK

This article is reproduced from the previous edition, volume 4, pp. 1787–1788, © 1994, Elsevier Ltd., with revisions made by the Editor.

Islam came to Southeast Asia peacefully and over several centuries. While mainland Southeast Asia by and large remained faithful to Buddhism in its various forms and expressions, Islam had a particular impact on the peoples of insular Southeast Asia speaking Western Austronesian or Indonesian languages. Islam thus took hold of the Indonesian archipelago, parts of the southern Philippines and of the Malay Peninsula, a region often referred to as 'the Malay world.' Southeast Asian Muslims generally follow Sunnite traditions and adhere to the Shafi'i school of law. There are doubts whether traces of an earlier presence of Shi'ite elements in West Sumatra indicate a significant change of allegiance or are due to fairly recent contacts with Indian Shi'ites brought to the region under the British at the beginning of the 19th century.

Although Islam found a more ready acceptance among (coastal) ethnic groups which had previously had contact with Indian religions, it did not, however, replace these completely. But, just as there are pockets of Hinduism and Buddhism alive in the archipelago and in particular on Java and Bali, so too Islam is to be found beyond the Peninsula – on the mainland in Burma, and among the Cham of Kampuchea and Vietnam who, significantly, speak an Indonesian language. The Cham, like so many Southeast Asian peoples, were converted to Islam through the Malays and are said to have received the faith from Brunei and Aceh. Yet, the importance of strong links with Kelantan in the north of the Peninsula, which appear to be of a more recent date, should not be underestimated. The origins of the small, indigenous Muslim community found in Burma must be sought entirely with Indian immigrants, and a once famous Muslim settlement in 17th-century Ayudhya was probably mainly populated by foreign merchants.

At the beginning of the 21st century, there are about 30 million Muslims living in Malaysia (the home of the great majority), Singapore, Brunei, Southern Thailand (Patani), and the Southern Philippines. Indonesia, with 170 million Muslims, is the country with the largest number of Muslim citizens in the world. There is no comprehensive theory to explain the process of Islamization of the Malay world, which followed a different course in almost every ethnic (and thus linguistic) group in which it occurred in the region. One assumes, however, that traders of various nationalities played an important part and that, at least in its early years, Islam came via the southern part of the Indian subcontinent, totally bypassing Moghul culture in the process. This is an important point, since Islam hardly influenced Southeast Asian sciences or culture in the broader sense, focusing almost exclusively on theology and law. As far as existing forms of art were concerned, it would appear that the acceptance of Islam led to the disappearance of many artistic and creative expressions which were not of an immediately religious nature. Others had to adapt radically in order to survive.

The Trengganu Stone inscription of the 13th century, which perhaps not surprisingly deals with legal matters, is the first example of Malay written in Arabic script and clear proof of the even earlier presence of the religion in the region, yet historical evidence would indicate that a more widely spread Islamization only began from the 15th century onwards. This was also the period when Sufi brotherhoods flourished internationally and when, after the Mongol conquest, Indian Ocean trade reached new peaks of intensity.

Once Islam had reached the archipelago and established a first foothold in the coastal trading cities of Sumatra, the Malay language became its main carrier, and it is not surprising that of the languages of Southeast Asia, it is Malay which has been influenced most by Islamic traditions. At some point, being a Malay even became synonymous with being a Muslim. Through Malay, the religion reached the northern coast of Java and most other islands. By the end of the 19th century, the process of Islamization of the Malay world can be said to have been completed, even though some Indonesian Muslims argue that after the conversion of the body, there still remains the task of converting the soul.

In the process of Islamization, Indian scripts and their derivatives were gradually replaced by a modified Arabic script which is closer to its Persian form. This is called *Jawi* in the Malay world where it is still used widely in preference to 'Roman' script, and *pegon* in Javanese (which otherwise adhered to its original, Indic script), where it is used exclusively for writing Islamic texts. Most significantly, *Jawi* provided for additional letters to convey the sounds of *ch, g, ng, ny,* and *p*, which are not found in standard Arabic.

Arabic also exerted its influence on syntax, morphology, and lexicography. While it is true to say that the influence on syntax and morphology is more clearly felt in the realm of theological literature, Qur'ānic exegesis, grammar, and jurisprudence, the influence of lexicography extended much further and

led to the reception and total absorption of a wide range of words of Arabic and Persian origin, extending well beyond the purely religious sphere. The important concept of traditional (non-Islamic) law and custom, which is seen in opposition to Islamic law, is only known under its Arabic term '*adat*.'

It is interesting, however, that even under the influence of Islam the Indic heritage was not completely suppressed, as is evinced by the survival of a Tamil word such as *lebai* for the mosque official and Sanskrit terms for key Islamic concepts such as the fast – *saum* in Arabic, but Sanskrit *puasa* in most Indonesian languages – and the religious disciple, who is still called a *santri*. Equally interesting is the fact that the planners of modern Indonesian, besides drawing on European languages, turn to Sanskrit rather than Arabic when it comes to the coining of new words and expressions. It is worth remembering, however, in this connection that in the late 20th century the influence on Indonesian of Javanese, which had always been more influenced by Sanskrit, was stronger than that of Malay. In view of the ethnic distribution of Indonesia's population, 60% of whom speak Javanese and only 5% Malay as their first language, this would appear natural, although non-Javanese Muslims find this Javanese influence hard to accept.

See also: Arabic; Austronesian Languages: Overview; Indonesia: Language Situation; Language Education Policies in Southeast Asia; Malayo-Polynesian Languages; Malaysia: Language Situation; Philippines: Language Situation; South Philippine Languages.

Bibliography

Bausani A (1975). 'Is classical Malay a 'Muslim language'?'. *Boletin de la Asociación Española de Orientalistas 11,* 111–121.

Jones R (ed.) (1978). *Indonesian etymological project III: Arabic Loan-words in Indonesian.* London: School of Oriental and African Studies.

Kratz E U (1990). 'Islam in Indonesia.' In Clarke P B (ed.) *The world's religions: Islam.* London: Routledge Reference.

Ronkel P S van (1899). 'Over invloed der Arabische syntaxis op de Maleische.' *Tijdschrift voor Indische Taal-, Land- en Volkenkunde 41,* 498–528.

Yegar M (1972). *The Muslims of Burma: a study of a minority group.* Wiesbaden: Otto Harrassowitz.

Islam in the Former Soviet Union

S Akiner, SOAS, University of London, UK; updated by **D Kurzon** University of Haifa, Haifa, Israel

Some 80 million Muslims (2003) live in the territories of the former Soviet Union, a Muslim population second in size only to those of such countries as Indonesia, Pakistan, Bangladesh, and India. Almost all are Sunni (as are the majority of Muslims in general) of the Hanafi school. There are, however, some Shi'ites in Azerbaijan and several thousand Ismailis, followers of the Aga Khan, in the Pamir mountains of Tadjikistan, in the regions adjacent to Pakistan, Afghanistan, and China.

Taken as a whole, the Muslims amount to some 20% of the total population of the old Soviet territories. However, there is considerable ethnic diversity among them. Some 85% are Turkic; of these, the Uzbeks (22.6 million) are the most numerous, constituting the third largest nationality after the Russians and the Ukrainians. Those belonging to the Iranian family account for 8% the Tadjiks (6 million) are by far the largest nationality in this group. The Caucasian family includes such peoples as the Avars, Laks, Lezgis, and Dargis; none very large in size, they speak languages that are almost wholly mutually incomprehensible.

Muslims are concentrated in three regions: in Central Asia (60%), Transcaucasia (20%), and the Volga region and Siberia (20%). Islam was brought to Central Asia and Transcaucasia by Arab invaders in the 7th century, scarcely 20 years after the death of the Prophet. It gradually gained ascendancy, to become in time the dominant force in those societies. The Tatars and Turkicized Mongols of the Golden Horde, who accepted Islam during the 13th and 14th centuries, carried the religion to the far side of the Urals, along the Volga and Siberia; it remained restricted to the tribes of the Golden Horde, however, and did not spread to the Slav population, who remained Christian.

The power of the Golden Horde began to wane during the 15th century, while that of Muscovy grew ever stronger. The capture of the Tatar strongholds of Kazan (1552) and Astrakhan (1556) by Ivan the Terrible brought Muslims under Russian control for the first time. For much of the next two centuries,

led to the reception and total absorption of a wide range of words of Arabic and Persian origin, extending well beyond the purely religious sphere. The important concept of traditional (non-Islamic) law and custom, which is seen in opposition to Islamic law, is only known under its Arabic term '*adat*.'

It is interesting, however, that even under the influence of Islam the Indic heritage was not completely suppressed, as is evinced by the survival of a Tamil word such as *lebai* for the mosque official and Sanskrit terms for key Islamic concepts such as the fast – *saum* in Arabic, but Sanskrit *puasa* in most Indonesian languages – and the religious disciple, who is still called a *santri*. Equally interesting is the fact that the planners of modern Indonesian, besides drawing on European languages, turn to Sanskrit rather than Arabic when it comes to the coining of new words and expressions. It is worth remembering, however, in this connection that in the late 20th century the influence on Indonesian of Javanese, which had always been more influenced by Sanskrit, was stronger than that of Malay. In view of the ethnic distribution of Indonesia's population, 60% of whom speak Javanese and only 5% Malay as their first language, this would appear natural, although non-Javanese Muslims find this Javanese influence hard to accept.

See also: Arabic; Austronesian Languages: Overview; Indonesia: Language Situation; Language Education Policies in Southeast Asia; Malayo-Polynesian Languages; Malaysia: Language Situation; Philippines: Language Situation; South Philippine Languages.

Bibliography

Bausani A (1975). 'Is classical Malay a 'Muslim language'?'. *Boletin de la Asociación Española de Orientalistas 11,* 111–121.

Jones R (ed.) (1978). *Indonesian etymological project III: Arabic Loan-words in Indonesian.* London: School of Oriental and African Studies.

Kratz E U (1990). 'Islam in Indonesia.' In Clarke P B (ed.) *The world's religions: Islam.* London: Routledge Reference.

Ronkel P S van (1899). 'Over invloed der Arabische syntaxis op de Maleische.' *Tijdschrift voor Indische Taal-, Land- en Volkenkunde 41,* 498–528.

Yegar M (1972). *The Muslims of Burma: a study of a minority group.* Wiesbaden: Otto Harrassowitz.

Islam in the Former Soviet Union

S Akiner, SOAS, University of London, UK; updated by **D Kurzon** University of Haifa, Haifa, Israel

Some 80 million Muslims (2003) live in the territories of the former Soviet Union, a Muslim population second in size only to those of such countries as Indonesia, Pakistan, Bangladesh, and India. Almost all are Sunni (as are the majority of Muslims in general) of the Hanafi school. There are, however, some Shi'ites in Azerbaijan and several thousand Ismailis, followers of the Aga Khan, in the Pamir mountains of Tadjikistan, in the regions adjacent to Pakistan, Afghanistan, and China.

Taken as a whole, the Muslims amount to some 20% of the total population of the old Soviet territories. However, there is considerable ethnic diversity among them. Some 85% are Turkic; of these, the Uzbeks (22.6 million) are the most numerous, constituting the third largest nationality after the Russians and the Ukrainians. Those belonging to the Iranian family account for 8% the Tadjiks (6 million) are by far the largest nationality in this group. The Caucasian family includes such peoples as the Avars, Laks, Lezgis, and Dargis; none very large in size, they speak languages that are almost wholly mutually incomprehensible.

Muslims are concentrated in three regions: in Central Asia (60%), Transcaucasia (20%), and the Volga region and Siberia (20%). Islam was brought to Central Asia and Transcaucasia by Arab invaders in the 7th century, scarcely 20 years after the death of the Prophet. It gradually gained ascendancy, to become in time the dominant force in those societies. The Tatars and Turkicized Mongols of the Golden Horde, who accepted Islam during the 13th and 14th centuries, carried the religion to the far side of the Urals, along the Volga and Siberia; it remained restricted to the tribes of the Golden Horde, however, and did not spread to the Slav population, who remained Christian.

The power of the Golden Horde began to wane during the 15th century, while that of Muscovy grew ever stronger. The capture of the Tatar strongholds of Kazan (1552) and Astrakhan (1556) by Ivan the Terrible brought Muslims under Russian control for the first time. For much of the next two centuries,

the Tatars suffered severe repression and religious persecution. It was not until the reign of Catherine the Great (1762–1796) that matters improved. During this period the community revived. A Muftiat was created in Orenburg in 1788 with responsibility for Muslim affairs. Tatars were granted a number of concessions in both religious and secular affairs; the latter included the restoration of the privileges of the Tatar nobility.

In the 19th century, many more Muslims came under Russian rule. Azerbaijan and the Caucasus were incorporated into the empire in the first half of the century, Central Asia in the second. For the most part, however, the Russian colonial administration interfered little in the everyday life of the indigenous peoples. Muslim institutions continued to function as before. Christian missionary activity was minimal.

When the Bolsheviks took power after the 1917 Revolution, one of their first acts was to reassure the "toiling Muslims of Russia and the East" that their beliefs and customs were henceforth "inviolable." Yet by the end of the 1920s, a fierce antireligious campaign had been launched throughout the USSR. In the Muslim regions, mosques were closed, the Qur'ān and other religious works confiscated, Muslim courts and schools abolished, teachers and scholars killed or imprisoned. The Arabic script, used for centuries by all the Muslims of Central Asia, Transcaucasia, Siberia, and the Volga, the visible symbol of Islam and symbolic link with fellow believers throughout the world, was replaced first by the Latin (1930), then the Cyrillic (1940). These changes of script effectively ensured that all forms of pre-Revolutionary literature, secular and religious alike, became inaccessible to future generations, because, quite simply, they became illegible.

During World War II, the Soviet government adopted a more conciliatory attitude toward religion. For the Muslims, four regional Spiritual Directorates were created, each headed by a Mufti, to administer such manifestations of Islam as were now allowed to reappear. The largest and most important Directorate was seated in Tashkent, the capital of Uzbekistan; its Mufti played the role of unofficial ambassador to Muslim communities abroad, acting as a mouthpiece for Soviet views on such matters as the Arab-Israeli conflict. Two *madrasa* (religious colleges) were reopened for the training of Muslim functionaries (in Tashkent and Bukhara), and public worship was permitted in a small number of mosques. Twenty to 30 carefully chosen representatives were allowed to go on the *hajj* (annual pilgrimage to Mecca); a tiny amount of religious literature was printed. Most of these measures were intended to impress Muslims abroad. They were too limited to have any real effect on the life of Soviet Muslims.

Truly fundamental changes did not come until the late 1980s, when the policy of *perestroika* ('restructuring') resulted in more mosques being opened in 1989 than had previously been allowed over several decades. Extensions to the two existing *madrasa* were sanctioned, and two new ones opened in Baku and Kazan. A new edition of the Qur'ān (50 000 copies) was printed in 1989, and translations into the local languages (Uzbek, Kazakh, etc.) were put in hand. A small but growing number of works on Islam are becoming available to the public at large. The Arabic script is being taught in schools (as an optional subject) and in evening classes; in some republics, there is a movement to have it reintroduced as the national script, in place of the Cyrillic script. In Tajikistan, for example, a law passed in the mid-1990s laid down that the Arabic script be used for writing the Tajik (Tajiki) language, but this law has never been implemented. What has happened is the replacement of the Cyrillic by the roman script. In some of the other republics, e.g., Kyrgyzstan, both Russian and the local language written in the Cyrillic script have been retained as official languages, while in others, e.g., Turkmenistan, Russian has been retained with the local language written in roman script as official languages.

The languages of Soviet Muslims underwent considerable change between the 1920s and the 1990s. Language planning policies were introduced in the 1920s. Consequently, new literary languages were consciously developed on the basis of selected dialects for many large and small ethnic groups. Some Muslim groups had had no standard, unified written form of their languages before this period; for others, such as the Uzbeks, the classical literary medium (Chagatai/Old Uzbek) was replaced by a form closer to the vernacular dialects. Large numbers of Russian loanwords (more accurately, Western European loanwords borrowed through Russian) were incorporated into these languages. In some cases, they introduced new concepts (e.g., in the fields of ideology, administration, Western-style science and technology); in others they replaced existing Arabic/Persian terms (e.g., in Uzbek, Arabic *dorilfunun* was replaced by Russian *universitet* 'university'). Occasionally, old and new forms coexisted as synonyms (e.g., in Uzbek, Arabic *iqtisod* and Russian *ekonomika* 'economy' are used as exact equivalents). Toward the end of the 1980s, a growing sense of resentment against the Russification of these languages became apparent. This heralded a move to reinstate the Arabic/Persian elements, parallel to the move to change the script from the Cyrillic back to the Arabic.

After the disintegration of the Soviet Union, the great problem that faces the leaders of Islam is how

to respond to the changed situation: in the past, their principal concern was to preserve the basic functional framework of the religion: now they are able, and indeed encouraged, to take a more active role in society, to help in its moral regeneration. It is not an easy task. They have yet to come to terms with it.

See also: Calligraphy, Islamic; Christianity in Central Asia and the Near East; Islam in the Near East; Kazakh; Tatar; Turkmen; Uzbek.

Bibliography

Akiner S (1986). *The Islamic peoples of the Soviet Union* (2nd edn.). [Contains some information on scripts and vocabulary.] London: Kegan Paul.

Sjoberg A F (1963). *Uzbek structural grammar.* Bloomington: Indiana University Press.

Waterson N (1980). *Uzbek–English dictionary.* Oxford: Oxford University Press.

Islam in the Near East

J N Mattock, University of Glasgow, Glasgow, UK

This article is reproduced from the previous edition, volume 4, pp. 1783–1785, © 1994, Elsevier Ltd., with revisions made by the Editor.

While Islam numbers among its adherents speakers of very many languages, it is dominated linguistically by Classical Arabic (*al-lug͟ha al-fuṣḥā* 'the pure language') to an extent that not even medieval Christianity can be said to have been dominated by Latin. The Qur'ān, as the actual word of God, must be read or recited in Arabic. It is widely disputed whether it is permissible to translate the Qur'ān and thus modify, if not pervert, the divine revelation. Many translations have, of course, been made, by Muslims as well as by non-Muslims, and an English translation was authorized in 1984 by the University of al-Azhar in Cairo, the most widely respected repository of Sunni orthodoxy, but the permissibility of translation remains a live issue. The prayers of the five canonical times must also be recited in Arabic. Only the Friday sermon is generally given in the vernacular in non-Arab communities.

Islam and Classical Arabic – A Symbiosis

It may reasonably be argued that it is the universal use of Arabic for religious purposes that largely sustains the very considerable cohesion of the Muslim brotherhood, bridging as it does the gulf between Sunni and Shi'i Islam, and even that between the Ahmadi movement and the rest. However imperfect a Muslim's command of Classical Arabic may be, it has become, as the medium in which the word of God was revealed, part of his heritage, which is to be revered and cherished.

It may also be argued that to an even greater degree, it is Islam that has, reciprocally, sustained Classical Arabic. The idiom of the Qur'ān is not precisely that of the other surviving literature of its period – the pre-Islamic and early Islamic poetry – or of that of later periods. This is inherent in the doctrine of the inimitability (*i'jāz*) of the Qur'ān. It is, however, closer to the language that has survived, with remarkably little change, to become what is now known as 'Modern Standard/Formal Arabic' than are, for example, most 16th-century forms of European languages to their modern counterparts. It is this language that the Arabs use for writing of all kinds, for formal oral communication – lectures, broadcasts, speeches, etc. – and, to some extent, for communication with speakers of 'dialects' that differ greatly from their own. It is the spoken native tongue of no one, and proficiency in its use varies greatly according to the degree of education of the user.

It is still uncertain what the relationship is between Classical Arabic and the various forms of speech (which are not generally represented in writing) that are familiarly referred to as the Arabic 'dialects,' collectively known in Arabic as *al-lug͟ha al-'āmmiyya* 'the common language,' *al-lug͟ha al-dārija* 'the current language,' or *al-lahajāt* 'the idioms, the dialects.' Traditionally, it has been claimed that these are simply offshoots of the classical language that have become differentiated from it and from one another in the course of time. However, the theory has been advanced that a more complicated process of pidginization and creolization among the non-Arab subject peoples of the Islamic empire should be recognized as having taken place. Whatever the answer, the result has been that these 'dialects' are as different from one another as are the various languages of the Romance group. They are, to all intents and purposes, separate languages and might well already have been

to respond to the changed situation: in the past, their principal concern was to preserve the basic functional framework of the religion: now they are able, and indeed encouraged, to take a more active role in society, to help in its moral regeneration. It is not an easy task. They have yet to come to terms with it.

See also: Calligraphy, Islamic; Christianity in Central Asia and the Near East; Islam in the Near East; Kazakh; Tatar; Turkmen; Uzbek.

Bibliography

Akiner S (1986). *The Islamic peoples of the Soviet Union* (2nd edn.). [Contains some information on scripts and vocabulary.] London: Kegan Paul.

Sjoberg A F (1963). *Uzbek structural grammar.* Bloomington: Indiana University Press.

Waterson N (1980). *Uzbek–English dictionary.* Oxford: Oxford University Press.

Islam in the Near East

J N Mattock, University of Glasgow, Glasgow, UK

This article is reproduced from the previous edition, volume 4, pp. 1783–1785, © 1994, Elsevier Ltd., with revisions made by the Editor.

While Islam numbers among its adherents speakers of very many languages, it is dominated linguistically by Classical Arabic (*al-lugha al-fuṣḥā* 'the pure language') to an extent that not even medieval Christianity can be said to have been dominated by Latin. The Qur'ān, as the actual word of God, must be read or recited in Arabic. It is widely disputed whether it is permissible to translate the Qur'ān and thus modify, if not pervert, the divine revelation. Many translations have, of course, been made, by Muslims as well as by non-Muslims, and an English translation was authorized in 1984 by the University of al-Azhar in Cairo, the most widely respected repository of Sunni orthodoxy, but the permissibility of translation remains a live issue. The prayers of the five canonical times must also be recited in Arabic. Only the Friday sermon is generally given in the vernacular in non-Arab communities.

Islam and Classical Arabic – A Symbiosis

It may reasonably be argued that it is the universal use of Arabic for religious purposes that largely sustains the very considerable cohesion of the Muslim brotherhood, bridging as it does the gulf between Sunni and Shi'i Islam, and even that between the Ahmadi movement and the rest. However imperfect a Muslim's command of Classical Arabic may be, it has become, as the medium in which the word of God was revealed, part of his heritage, which is to be revered and cherished.

It may also be argued that to an even greater degree, it is Islam that has, reciprocally, sustained Classical Arabic. The idiom of the Qur'ān is not precisely that of the other surviving literature of its period – the pre-Islamic and early Islamic poetry – or of that of later periods. This is inherent in the doctrine of the inimitability (*i'jāz*) of the Qur'ān. It is, however, closer to the language that has survived, with remarkably little change, to become what is now known as 'Modern Standard/Formal Arabic' than are, for example, most 16th-century forms of European languages to their modern counterparts. It is this language that the Arabs use for writing of all kinds, for formal oral communication – lectures, broadcasts, speeches, etc. – and, to some extent, for communication with speakers of 'dialects' that differ greatly from their own. It is the spoken native tongue of no one, and proficiency in its use varies greatly according to the degree of education of the user.

It is still uncertain what the relationship is between Classical Arabic and the various forms of speech (which are not generally represented in writing) that are familiarly referred to as the Arabic 'dialects,' collectively known in Arabic as *al-lugha al-'āmmiyya* 'the common language,' *al-lugha al-dārija* 'the current language,' or *al-lahajāt* 'the idioms, the dialects.' Traditionally, it has been claimed that these are simply offshoots of the classical language that have become differentiated from it and from one another in the course of time. However, the theory has been advanced that a more complicated process of pidginization and creolization among the non-Arab subject peoples of the Islamic empire should be recognized as having taken place. Whatever the answer, the result has been that these 'dialects' are as different from one another as are the various languages of the Romance group. They are, to all intents and purposes, separate languages and might well already have been

recognized as such but for two factors, which are partly interconnected.

The first of these factors is Arab nationalism. This, although fostered by those who have had an interest in so doing, has never proved particularly powerful and is indeed somewhat specious; the great majority of those who consider themselves Arabs have little enough direct link with the original Arab conquerors of the 7th century. The second is the importance that the Arabic language has for Islam. It is this that has, on the one hand, preserved it virtually in its pristine form and, on the other, given it a status that has caused its offspring to be regarded, even by those who speak them, as inferior. The 'dialects' are referred to in terms of greater or lesser 'purity' according to the degree to which they diverge from the classical norm. That this is attributable to the influence of Islam is suggested by the case of Maltese; here one finds a descendant of Arabic that has developed in an entirely non-Muslim environment over more than seven centuries, has adopted its own Romanized system of writing, and has been recognized as a language in its own right. Elsewhere, non-Muslim speakers of various 'dialects' have formed minorities within territories under Muslim domination.

Classical Arabic, then, is in effect a hieratic language for Muslims, although it is hardly recognized as such by those who consider themselves Arabic speaking. The extent to which the vernaculars are permeated with classical expressions, largely of a religious nature, perhaps tends to disguise the fact. Thus, the frequent citation of verses from the Qur'ān and the Hadith of the Prophet, exclamations such as *al-ḥamdu li-Lāh* 'praise be to God!', *bismi l-Lāhi r-Raḥmāni r-Raḥīm* 'in the name of God, the Compassionate, the Merciful', *a'ūdhu bi-l-Lāh* 'I take refuge in God (i.e., God forbid!)', and even the customary exchange of greetings, *as-salāmu 'alaykum* 'peace be upon you!' – *wa-'alaykumu s-salām* 'and upon you be peace!', in all of which the inflections of the classical language, absent from the vernaculars, are generally retained, strengthen the belief that other features of everyday speech represent a falling short from a classical perfection. On a secular level, the predilection of the educated for quoting medieval classical poetry, in which the inflections are necessary for the meter, also reinforces this. Thus, although Classical Arabic is not a natural medium for them, 'Arabic-speaking' Muslims do not feel it to be so divorced from their own native speech as, for example, Russian speakers regard Old Church Slavonic. For non-Arab Muslims, of course, Arabic is a hieratic language. Since, however, in many cases their familiarity with it is exclusively in a religious context – whether simply as used in worship or also as encountered in the study of secondary theological literature – they may not be in a position to differentiate between 'varieties' of 'Arabic.' In addition, they are disposed to follow the Arabs in regarding 'Arabic' as one language, proficiency in which is attained by different people to different degrees. Thus, although it is to them a hieratic language, it is not one that is equally remote from all believers; their Arab coreligionists have the enviable advantage of possessing it as their native medium of communication.

Other Islamic Languages

There is no documentation extant for the evolution of Pahlavi into Islamic Persian during the first three centuries after the Arab conquest; all Persians apparently wrote in Arabic, even though they presumably still spoke some form of their native tongue. When Persian emerged, it was heavily arabicized. A large proportion of its vocabulary was borrowed directly, and Arabic phrases were also adopted wholesale – not merely pious expressions and other religious idioms, but ordinary secular elements as well. It was at first still possible to use a less arabicized version, as can be seen in Firdawsi's epic the *Shāhnāmah*, in which, since it is concerned with preconquest history, Arabic phraseology would clearly be inappropriate. This, however, appears to have been regarded as deliberate archaizing, and although the *Shāhnāmah* has always remained immensely popular, no subsequent writers followed Firdawsi's example. The degree to which the language of individual authors was consciously arabicized, however, differs greatly; in general, prose is more affected than verse, and the more orotund the style adopted, the more Arabic it appears.

Oddly enough, some items of Persian religious terminology survive, used interchangeably with their Arabic equivalents, and sometimes in preference to these, for example, *namāz* 'prayer' (Arabic *salāt*), *payghāmbar* 'messenger, prophet' (Arabic *rasūl, nabī*), and *Khudā* 'God' (Arabic *Allāh*). Arabic, however, was a dominant influence; sometimes even Persian syntax seems to be approximating to Arabic. Persian and Arabic words were readily compounded, as in *khidmatgar* 'servant' (Arabic *khidma* 'service' + *-gar*, a Persian suffix indicating 'one who does something'), and *dawlat-bar-andāz* 'revolutionary' (Arabic *dawla* 'state' + *bar-andāz*, a participial form from the Persian verb *bar-andākhtan* 'to cast down'); Persian plural endings were attached to Arabic singular nouns, side by side with Arabic plurals, e.g., *ḥālathā* (Persian plural suffix *-hā*, for inanimate objects) and *ḥālāt* as plurals of *ḥāla* 'state/condition' and *ghulāmān* (Persian plural suffix *-ān*, generally for animate beings) and *ghilmān* as plurals of *ghulām* 'boy/slave'.

It appears possible that the Persian word *musulmān* 'Muslim' may have come into existence through a misreading of such a hybrid, *muslimān* 'Muslims' (*muslim* + plural suffix -*ān*), requiring the formation of a further plural *musulmānān*.

The languages that appeared somewhat later on the Islamic scene, such as Turkish and Urdu, borrowed largely from Persian, taking over in the process the Arabic borrowings already naturalized there. Persian itself, from about the 15th century, came to replace Arabic as the principal cultural, if not religious, language of Eastern Islam. It continued to enjoy this position until the last century; of the poetry of Muhammad Iqbal, the celebrated Urdu poet, a sizable proportion is in Persian.

Attempts have been made to 'purify' both Turkish and Persian by replacing the 'alien' elements with 'native' ones. Atatürk's efforts in the 1920s to do this for Turkish had some success; modern Turkish contains a great deal less Arabic and Persian vocabulary than Ottoman Turkish, although much remains, some of it, no doubt, unrecognized as such. The last Shah, Muhammad Reza, became an enthusiastic propagandist for the 'Aryan' character of Persian civilization, and accordingly encouraged the resurrection of Persian terms to replace Arabic; thus, 'agriculture' became *kashāvarzī* instead of *zirā'at* and 'student' *dānish-jū* (literally, 'knowledge seeker') instead of *ṭālib*. How much of this linguistic engineering will survive the Islamic revolution, with its inevitable emphasis on Arabic, remains to be seen. In fact, both the Turkish and the Iranian initiatives may be regarded as moves in campaigns of secularization. In Atatürk's case, the campaign was overt and acknowledged; in Muhammad Reza's case, it was covert, but nevertheless real, consequent upon the opposition that both his and his father's policies of modernization had encountered from the Iranian religious leaders.

The Arabic Script

The Arabic script, with various modifications, was almost universally adopted for other languages of the Muslim world; this remained the case until earlier this century, constituting a further bond between Muslims. The respect in which it was held amounted almost to reverence. The art of calligraphy was highly valued, and the Arabic script, often in highly ornate and fantastic forms, became one of the principal features of all the decorative arts. Poetry abounds in metaphors and similes taken from the shapes of the letters.

It has been retained for Persian, Urdu, and Pashto, all three of which are, of course, the languages of regions in which Islam is particularly strong. With some qualifications, it is reasonably well suited to these languages, although the absence of short vowels means that they cannot be read aloud by someone who does not know them. Turkish abandoned the Arabic script in 1928. It was not well suited to a language with a sophisticated vocalic system, but this reform was also a further step in Atatürk's attempt to diminish the influence of Islam in Turkey; it distanced the newly literate both from the Qur'ān and associated Arabic writing and from the corpus of Islamic Turkish and Persian literature. The modified Latin orthography of modern Turkish works well, but it does tend to disguise the Arabic and Persian borrowings in its reproduction of their Turkish pronunciation: the elements *ajzā'* 'components, drugs' (plural of Arab *juz'*) and *khānah* 'house, shop' (Persian) are barely recognizable in *eczane* 'chemist's shop'; *fevkalâde* 'extraordinary' appears somewhat far removed from its Arabic original *fawq al-'āda*. Kurdish, which is, somewhat distantly, related to Persian, has modified the Arabic script in such a way that short vowels can be represented; this expedient also disguises the shape of the Arabic, Persian, and Turkish elements, but it appears to have had no antireligious origin.

See also: Arabic; Islam and Arabic; Maltese; Qur'an.

Bibliography

Fārūqī I R & Fārūqī L L (1986). *The cultural atlas of Islam.* New York: Macmillan.

Gibb H A R *et al.* (eds.) (1960). *The encyclopedia of Islam.* Leiden: E. J. Brill. [esp. articles 'Arabiyya, Islām, Lugha.]

Holt P M, Lambton A K S & Lewis B (eds.) (1970). *The Cambridge history of Islam.* Cambridge: Cambridge University Press.

Versteegh K (1984). *Pidginization and creolization: the case of Arabic. Amsterdam Studies in the Theory and History of Linguistic Science,* series 4, (vol. 33). Amsterdam: Benjamins.

Island Constraints

A Carnie, University of Arizona, Tucson, AZ, USA

The Phenomena

Chomsky (1964) observed that there was a restriction on the family of displacement operations collectively known as A-bar movement or extraction (including topicalization, wh-movement, relative clause formation, and others), such that given two noun phrases (NPs), one embedded inside the other, only the topmost NP can be displaced. Ross (1967) suggested that this phenomenon is part of a larger pattern of constraints on movement that limit movement from certain kinds of constituents. In Ross's terminology, a syntactic constituent that disallows movement from within itself is known as an 'island.' The constraints that govern this behavior are known as 'island constraints.'

As a point of contrast, consider the following pair of grammatical sentences, where the second is taken to be a question version of the first:

(1a) I said [$_{\text{CP}}$ that John wanted [$_{\text{TP}}$ Mary to ask Bill [$_{\text{TP}}$ to bake **a cake**.]]]
(1b) **What** did you say that John wanted Mary to ask Bill to bake *t*?

Movement of the wh-word (*what*) that is the object of the verb *bake* is allowed to cross a large amount of structure. Now consider movement out of a similar sentence, but one where the starting place of the wh-word is in a clause that is inside an NP. Sentence (2) shows an example where the embedded clause is a complement to the noun; sentence (3) shows an example where the embedded structure is a relative clause:

(2a) I concealed [$_{\text{NP}}$ the fact [$_{\text{CP}}$ that John baked **a cake**].
(2b) ??**What** did you conceal the fact that John baked *t*?

(3a) I saw [$_{\text{NP}}$ the man [$_{\text{CP}}$ who baked **that cake**]]
(3b) ***Which cake** did you see the man who baked ***t***?

The sentences in (2b) and (3b) are subject to what Ross called the '*Complex NP Constraint*' (CNPC), and are known as (complex) NP-islands. This constraint simply says that extraction out of a clause within an NP is ungrammatical.

NPs are not the only island, however; Ross observes that any element (including clauses) in subject position is also an island for extraction. This can be seen in (4); (4a) and (4b) show a pair where we are extracting an object out of a tensed clause in subject position. (4c) and (4d) show the same thing with a nonfinite subject clause.

(4a) [$_{\text{CP}}$ That the police would arrest **several rioters**] was a certainty.
(4b) ***What** was that the police would arrest *t* a certainty?
(4c) [$_{\text{TP}}$ To read ***Great Expectations*** in peace] was all he wanted.
(4d) *What was to read ***t*** in peace all he wanted?

This is Ross's *Sentential Subject Constraint* (SSC).

Closely related to the CNPC and SSC is Ross's *Left Branch Condition* (LBC), which restricts movement of an NP that is the left branch of a larger NP.

(5a) I liked [$_{\text{NP}}$ [$_{\text{NP2}}$ **John**]'s friend].
(5b) ***Whose** did you like ***t*** friend?

Finally, Ross also proposed the *Coordinate Structure Constraint* (CSC), which disallows extraction of either member of a conjunction:

(6a) I liked **Mary** and **John**.
(6b) *Who did you like Mary and ***t***?
(6c) *Who did you like ***t*** and John?

In addition to Ross's islands, a number of other island types have been proposed, including negative islands (Schafer, 1995) and wh-islands (Chomsky, 1973). The latter are perhaps the most important kind of island. Wh-islands involve displacement of two wh-elements. One movement is to the front of the embedded clause; the second crosses over the first, and puts the wh-element at the beginning of some higher clause. (7a) represents a declarative clause. (7b) and (7c) show that movements of both the embedded NPs are acceptable independent of one another. (7d) shows that movement of both NPs together results in ungrammaticality.

(7a) Bill said [$_{\text{CP}}$ **Mary** loves **pineapples**.]
(7b) Bill asked [$_{\text{CP}}$ **what**$_i$ Mary loves t_i.]
(7c) **Who**$_j$ did Bill say [$_{\text{CP}}$ t_j loves pineapples.]
(7d) *Who$_j$ did Bill ask [$_{\text{CP}}$ what$_i$ t_j loves t_i]

The Subjacency Condition

In the government and binding theory of the late 1970s and 1980s, an attempt was made to capture all of the island phenomena (or a significant portion thereof) with one formal constraint. The most successful attempt at this came in the form of the *Subjacency Condition* (Chomsky, 1973; Koster, 1978; Rizzi, 1982); this constraint limited wh-movement by specifying dominating categories as ***bounding nodes***. Crossing too many bounding nodes in one swoop

resulted in ungrammaticality. A typical formalization of the subjacency condition is given in (8):

(8) A-bar movement may not cross two or more bounding nodes. Bounding nodes are NP and TP (= S).

In order to see how this accounts for islands, it is worth showing how it allows for long distance wh-movement as in (1). The underlying idea is that movement is *cyclic*. Although the movement in (1) appears to be long distance, it actually happens in shorter hops. Each hop lands in the specifier of a CP. As such, most cases of movement will only cross one bounding node (a single TP). This can be seen in (9) below, where each bounding node is circled. Each step in the movement crosses only one bounding node.

(9a) [$_{CP}$ did [$_{TP}$ you think [$_{CP}$ what$_i$ that [$_{TP}$ Millie said t_i]]]]?

move #1

(9b) [$_{CP}$ What did [$_{TP}$ you think [$_{CP}$ t_i that [$_{TP}$ Millie said t_i]]]]?

move #2

Contrast this with an NP-island, such as (2b), the structure of which is given in (10):

(10) ?? [$_{CP}$ **What** did [$_{TP}$ you conceal [$_{NP}$ the fact [$_{CP}$ that [$_{TP}$ John baked *t*]]]]]?

Successive cyclic movement will allow the first movement of the wh-word to cross only a single bounding node, which is acceptable:

(11) ?? [$_{CP}$ did [$_{TP}$ you conceal [$_{NP}$ the fact [$_{CP}$ what that [$_{TP}$ John baked t_i]]]]]?

move #1

However, subsequent movement to the highest CP specifier will result in a Subjacency violation, crossing both the NP node and the higher of the two TP nodes:

(12) ?? [$_{CP}$ What did [$_{TP}$ you conceal [$_{NP}$ the fact [$_{CP}$ *t* that [$_{TP}$ John baked t_i]]]]]?

**move #2*

A similar account can be given to wh-islands: movement of the first wh-word does not violate the constraint:

(13) *[$_{CP}$ did [$_{TP}$ Bill ask [$_{CP}$ what$_i$ [$_{TP}$ who loves t_i]]]]

move #1

Movement of the other wh-word will necessarily result in a Subjacency violation, because it will not be allowed to stop in the embedded specifier of the CP (which is occupied by *what*), so it will cross two TP bounding nodes:

(14) *[$_{CP}$ who$_j$ did [$_{TP}$ Bill ask [$_{CP}$ what$_i$ [$_{TP}$ t_j loves t_i]]]]

**move #2*

The subjacency constraint also explains the left branch condition, the sentential subject constraint, and some cases of the coordinate structure constraint (although not all). An important outgrowth of the research into subjacency was the observation that there was crosslinguistic variation in what counted as a bounding node. For example, Rizzi (1982) observed that CPs (S′) and not TPs (S) in Italian were bounding nodes. This kind of observation was influential in the development of casting crosslinguistic variation in terms of narrow parameters.

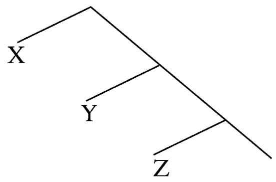
The Empty Category Principle and the Condition on Extraction Domains

Huang (1982) observed that extraction from adjuncts and specifiers was consistently worse than that of extraction of complements. Huang proposed the *Condition on Extraction Domains* (CED), which limits grammatical extraction to positions that are governed. (Roughly, government corresponds locally c-commanded by a lexical [theta role-assigning] head or by an antecedent.) The heart of this idea is contained in Chomsky's (1981) *Empty Category Principle* (ECP), which requires that the traces of A-bar movement be governed. Chomsky (1986) attempted to unify the ECP, subjacency, and other locality conditions (such as the binding conditions) in a system called *Barriers*. At the time of the writing of this entry, Barriers is widely considered to have been a failure.

Relativized Minimality

Rizzi (1990) first observed that every kind of movement (A-bar movement, A-movement (or NP-movement) and head movement) has some kind of locality constraint on the relationship between the displaced item and its starting position or trace. In each case, the moved item must c-command the position it moves to; and there can be no potential intervening candidate for the element at the top of the movement chain. In other words, given a structure such as (15), where Y and X are of a like type of category (wh-element, or NP or head), then a movement chain formed between X and Z is ungrammatical.

(15)



This differs from previous approaches where what was at stake in locality conditions was the nature of the categories that dominate the trace. Here, what is of relevance is the nature of the categories and positions that c-command the trace.

Island Constraints in the Minimalist Program

The relativized minimality approach has been very influential in the formulation of constraints on island phenomena in the minimalist program. In the earliest versions of the program (e.g., Chomsky, 1994), island effects were captured by making reference to a version of relativized minimality, whereby there are economy constraints that compare derivations of sentences and look for the most economical. One such constraint is *Shortest Move* (SM). This constraint gives preference to derivations where each movement moves to the closest potential landing site. Skipping that landing site (for example, because it is filled with another wh-word) will result in a dispreferred derivation. In later versions of the program (e.g., Chomsky, 1995), SM is replaced by an absolute condition that is part of the definition of the movement operation (known as ATTRACT). This condition is the *Minimal Link Condition*, which for any given landing site always prefers the closest potential filler. Refinements to this approach can be found in Collins (1997), and criticisms can be found in Johnson and Lappin (1999). In the most recent version of the minimalist program (Chomsky, 2000), known as *phase theory*, island effects reduce to the way in which the syntactic tree is derived. Very small, local structures (known as *phases*) are generated piece by piece from the bottom of the tree. Operations may only apply within these phases, or from the topmost position within the phase (the phase edge); this ensures that each step in the movement will be limited to occur either within a given phase or between adjacent phases, giving rise to successive cyclicity.

Approaches to Island Effects in Other Theories

Outside of Chomskyan linguistics, other approaches to syntax have attempted to capture some of the same results. Following the work of Gerald Gazdar (1981) in generalized phrase structure grammar, proponents of head-driven phrase structure grammar (HPSG) make use of the fact that the theory relies on licensing constraints on subtree structures (similar in many ways to the licensing constraints that hold of minimalist phasing). In HPSG, island conditions are at least partly stated as conditions on these licensing constraints. For example, the CSC falls out from the requirement that the rule which licenses coordination stipulates that neither conjunct may contain a gap (see Sag *et al.* [2004] for a textbook treatment, and Johnson and Lappin [1999] for a more in-depth analysis). In lexical functional grammar, long-distance dependencies are expressed in terms of the notion of functional uncertainty, whereby functional relations such as FOCUS (which governs wh-questions) search for a potential variable to satisfy their features. There are conditions on the search, called inside-out constraints, which limit the search space within the f-structure (see Falk, 2001 for an accessible introduction to these ideas). In the functionalist literature, the most common approach is to try to explain subjacency as a processing or memory constraint. For example, Pritchett (1991) suggests that subjacency reduces to the desire of the parser to match up displaced elements and gaps. When too many choices are available, the parser crashes. For a critical review of such approaches, see Lasnik (1999).

See also: Chomsky, Noam (b. 1928); Complement Clauses; Constituent Structure; Generative Grammar; Principles and Parameters Framework of Generative Grammar; Relative Clauses; Syntactic Variables and Variable-free Syntax; Syntactic Variation.

Bibliography

Chomsky N (1964). *Current issues in linguistic theory.* The Hague: Mouton.

Chomsky N (1973). 'Conditions on transformations.' In Anderson S & Kiparksy P (eds.) *A festschrift for Morris Halle.* New York: Holt, Rinehart & Winston. 232–286.

Chomsky N (1981). *Lectures on government and binding.* Dordrecht: Foris.

Chomsky N (1986). *Barriers.* Cambridge, MA: MIT Press.

Chomsky N (1994). 'A minimalist program for linguistic theory.' In Hale K & Keyser S J (eds.) *The view from Building 20.* Cambridge, MA: MIT Press.

Chomsky N (1995). *The minimalist program.* Cambridge, MA: MIT Press.

Chomsky N (2000). 'Minimalist inquiries: the framework.' In Martin R, Michaels D & Uriagereka J (eds.) *Step by step: essays on minimalist syntax in honor of Howard Lasnik.* Cambridge, MA: MIT Press. 89–156.

Collins C (1997). *Local economy.* Cambridge, MA: MIT Press.
Falk T (2001). *Lexical-functional grammar.* Stanford, CA: CSLI.
Gazdar G (1981). 'Unbounded dependencies and coordinate structure.' *Linguistics Inquiry 12,* 155–184.
Huang C-T J (1982). 'Logical relations in Chinese and the theory of grammar.' Ph.D. dissertation, MIT.
Johnson D & Lappin S (1999). *Local constraints vs. economy.* Stanford, CA: CSLI.
Koster J (1978). *Locality principles in syntax*. Dordrecht: Foris.
Lasnik H (1999). 'On the locality of movement.' In Darnell M, Moravscik E, Noonan M, Newmeyer F J & Wheatley K M (eds.) *Functionalism and formalism in linguistics 1: General papers*. Philadelphia: John Benjamins. 33–54.
Pritchett B (1991). 'Subjacency in a principle-based parser.' In Berwick R C, Abney S P & Tenny C (eds.) *Principle-based parsing: computation and psycholinguistics*. Dordrecht: Kluwer Academic. 301–345.
Rizzi L (1982). *Issues in Italian syntax*. Dordrecht: Foris.
Rizzi L (1990). *Relativized minimality.* Cambridge, MA: MIT Press.
Ross J R (1967). 'Constraints on variables in syntax.' Doctoral dissertation, MIT.
Sag I, Wasow T & Bender E (2004). *Syntax: a formal introduction.* Palo Alto, CA: Stanford University Press.
Schafer R (1995). 'Negation and verb second in Breton.' *Natural Language and Linguistic Theory 13,* 135–172.

Isle of Man: Language Situation

Editorial Team

Official languages on this self-governing crown dependency (estimated inhabitants: 77 655) of the United Kingdom are English and Manx. Manx Gaelic, a Goidelic language, was widely used earlier, but became extinct with the death of its last speaker in 1974. Thanks to a scholarly revival by enthusiasts, Manx has been formally taught in schools since 1992 and is now the second language of several hundred people. Government departments use English and Manx titles on all official documents, correspondence, and vehicles, and Manx is also used on town and road signs and increasingly by commercial businesses.

See also: United Kingdom: Language Situation.

Isolated Language Varieties

C Dannenberg, Virginia Tech, Blacksburg, VA, USA

The term 'isolation' and its corollary 'mobility' have continued to be key factors in the determination of the actuation, embedding, and trajectory of linguistic change. Although isolation has rarely been studied in and of itself as an independent linguistic variable (Chambers, 1995), dialectologists have been irresistibly drawn to isolated communities (and to isolated speakers for that matter) for their import descriptively and theoretically, since Gauchat's (1905) cornerstone study of the community of Charmey in the Swiss Alps. Such communities yield vital evidence into earlier stages of various dialects, including African American Vernacular English, Appalachian English, and even mainstream American English, to name a few. Additionally, insular and postinsular communities allow for rigorous investigation into language variation and change, particularly with respect to those language communities that are moribund – that is, those that are in danger of dying out within a generation or two. The case of Faetar, a Franco-Provençal dialect spoken in the villages of Gaeto and Celle in southern Italy and historically brought to the region by French immigrants in the 14th century, is such an example of a moribund language variety. Analysis of the phonological system of Faetar, particularly the phonological status of the schwa (absent in standard Italian), has provided rich data on the specifics of how language contact affects each of the languages in the contact situation (Nagy, 2001).

Quite recently, though, the concept of isolation and the assumed affects of isolation on dialect

Collins C (1997). *Local economy.* Cambridge, MA: MIT Press.
Falk T (2001). *Lexical-functional grammar.* Stanford, CA: CSLI.
Gazdar G (1981). 'Unbounded dependencies and coordinate structure.' *Linguistics Inquiry 12,* 155–184.
Huang C-T J (1982). 'Logical relations in Chinese and the theory of grammar.' Ph.D. dissertation, MIT.
Johnson D & Lappin S (1999). *Local constraints vs. economy.* Stanford, CA: CSLI.
Koster J (1978). *Locality principles in syntax.* Dordrecht: Foris.
Lasnik H (1999). 'On the locality of movement.' In Darnell M, Moravscik E, Noonan M, Newmeyer F J & Wheatley K M (eds.) *Functionalism and formalism in linguistics 1: General papers.* Philadelphia: John Benjamins. 33–54.
Pritchett B (1991). 'Subjacency in a principle-based parser.' In Berwick R C, Abney S P & Tenny C (eds.) *Principle-based parsing: computation and psycholinguistics.* Dordrecht: Kluwer Academic. 301–345.
Rizzi L (1982). *Issues in Italian syntax.* Dordrecht: Foris.
Rizzi L (1990). *Relativized minimality.* Cambridge, MA: MIT Press.
Ross J R (1967). 'Constraints on variables in syntax.' Doctoral dissertation, MIT.
Sag I, Wasow T & Bender E (2004). *Syntax: a formal introduction.* Palo Alto, CA: Stanford University Press.
Schafer R (1995). 'Negation and verb second in Breton.' *Natural Language and Linguistic Theory 13,* 135–172.

Isle of Man: Language Situation

Editorial Team

Official languages on this self-governing crown dependency (estimated inhabitants: 77 655) of the United Kingdom are English and Manx. Manx Gaelic, a Goidelic language, was widely used earlier, but became extinct with the death of its last speaker in 1974. Thanks to a scholarly revival by enthusiasts, Manx has been formally taught in schools since 1992 and is now the second language of several hundred people. Government departments use English and Manx titles on all official documents, correspondence, and vehicles, and Manx is also used on town and road signs and increasingly by commercial businesses.

See also: United Kingdom: Language Situation.

Isolated Language Varieties

C Dannenberg, Virginia Tech, Blacksburg, VA, USA

The term 'isolation' and its corollary 'mobility' have continued to be key factors in the determination of the actuation, embedding, and trajectory of linguistic change. Although isolation has rarely been studied in and of itself as an independent linguistic variable (Chambers, 1995), dialectologists have been irresistibly drawn to isolated communities (and to isolated speakers for that matter) for their import descriptively and theoretically, since Gauchat's (1905) cornerstone study of the community of Charmey in the Swiss Alps. Such communities yield vital evidence into earlier stages of various dialects, including African American Vernacular English, Appalachian English, and even mainstream American English, to name a few. Additionally, insular and postinsular communities allow for rigorous investigation into language variation and change, particularly with respect to those language communities that are moribund – that is, those that are in danger of dying out within a generation or two. The case of Faetar, a Franco-Provençal dialect spoken in the villages of Gaeto and Celle in southern Italy and historically brought to the region by French immigrants in the 14th century, is such an example of a moribund language variety. Analysis of the phonological system of Faetar, particularly the phonological status of the schwa (absent in standard Italian), has provided rich data on the specifics of how language contact affects each of the languages in the contact situation (Nagy, 2001).

Quite recently, though, the concept of isolation and the assumed affects of isolation on dialect

Collins C (1997). *Local economy.* Cambridge, MA: MIT Press.
Falk T (2001). *Lexical-functional grammar.* Stanford, CA: CSLI.
Gazdar G (1981). 'Unbounded dependencies and coordinate structure.' *Linguistics Inquiry 12,* 155–184.
Huang C-T J (1982). 'Logical relations in Chinese and the theory of grammar.' Ph.D. dissertation, MIT.
Johnson D & Lappin S (1999). *Local constraints vs. economy.* Stanford, CA: CSLI.
Koster J (1978). *Locality principles in syntax*. Dordrecht: Foris.
Lasnik H (1999). 'On the locality of movement.' In Darnell M, Moravscik E, Noonan M, Newmeyer F J & Wheatley K M (eds.) *Functionalism and formalism in linguistics 1: General papers*. Philadelphia: John Benjamins. 33–54.
Pritchett B (1991). 'Subjacency in a principle-based parser.' In Berwick R C, Abney S P & Tenny C (eds.) *Principle-based parsing: computation and psycholinguistics*. Dordrecht: Kluwer Academic. 301–345.
Rizzi L (1982). *Issues in Italian syntax*. Dordrecht: Foris.
Rizzi L (1990). *Relativized minimality.* Cambridge, MA: MIT Press.
Ross J R (1967). 'Constraints on variables in syntax.' Doctoral dissertation, MIT.
Sag I, Wasow T & Bender E (2004). *Syntax: a formal introduction.* Palo Alto, CA: Stanford University Press.
Schafer R (1995). 'Negation and verb second in Breton.' *Natural Language and Linguistic Theory 13,* 135–172.

Isle of Man: Language Situation

Editorial Team

Official languages on this self-governing crown dependency (estimated inhabitants: 77 655) of the United Kingdom are English and Manx. Manx Gaelic, a Goidelic language, was widely used earlier, but became extinct with the death of its last speaker in 1974. Thanks to a scholarly revival by enthusiasts, Manx has been formally taught in schools since 1992 and is now the second language of several hundred people. Government departments use English and Manx titles on all official documents, correspondence, and vehicles, and Manx is also used on town and road signs and increasingly by commercial businesses.

See also: United Kingdom: Language Situation.

Isolated Language Varieties

C Dannenberg, Virginia Tech, Blacksburg, VA, USA

The term 'isolation' and its corollary 'mobility' have continued to be key factors in the determination of the actuation, embedding, and trajectory of linguistic change. Although isolation has rarely been studied in and of itself as an independent linguistic variable (Chambers, 1995), dialectologists have been irresistibly drawn to isolated communities (and to isolated speakers for that matter) for their import descriptively and theoretically, since Gauchat's (1905) cornerstone study of the community of Charmey in the Swiss Alps. Such communities yield vital evidence into earlier stages of various dialects, including African American Vernacular English, Appalachian English, and even mainstream American English, to name a few. Additionally, insular and postinsular communities allow for rigorous investigation into language variation and change, particularly with respect to those language communities that are moribund – that is, those that are in danger of dying out within a generation or two. The case of Faetar, a Franco-Provençal dialect spoken in the villages of Gaeto and Celle in southern Italy and historically brought to the region by French immigrants in the 14th century, is such an example of a moribund language variety. Analysis of the phonological system of Faetar, particularly the phonological status of the schwa (absent in standard Italian), has provided rich data on the specifics of how language contact affects each of the languages in the contact situation (Nagy, 2001).

Quite recently, though, the concept of isolation and the assumed affects of isolation on dialect

maintenance and change have come under considerable scrutiny, and this criticism has led to dynamic definitions of isolation itself as well as the formulation of comprehensive principles by which to measure the influence of insular and postinsular communities on language change.

The Concept of Isolation

Categories of Isolation

Sociolinguists use the term 'isolation' to indicate relative degree of separation of a particular language community from surrounding contact varieties. Categories of separation may include (Montgomery 2000: 44–45).

- Physical (involving proximity to other communities, especially towns)
- Sociological (involving the frequency and variety of contact with other communities)
- Economic (involving the exchange of goods, services, and ideas with other communities)
- Psychological (involving the orientation and affinity of one community toward others, attachments to one's own culture, and openness to change)
- Cultural (involving the maintenance of distinctive practices and beliefs)
- Technological

Crucially, these categories are not discrete. In fact, each one may be dependent on, connected with, or reinforced by another. For example, sociological isolation, where speakers are separated in their day-to-day contact from surrounding communities, may be directly predicated by the community's physical remoteness. An example is the community of Smith Island, Maryland, where the islanders are quite socially isolated. Smith Island is not a tourist spot or an ideal place for retirement. Moreover, the island does not attract newcomers who are looking for a place to settle. In fact, the population of Smith Island is declining quickly, down from 650 in 1970 to 350 today. Certainly, the community's physical remoteness has played a role in this social isolation, as it is not and more than likely will never be accessible by bridge from the mainland, and the island is only accessible by day via private ferry service (see more discussion of Smith Island below).

Additionally, a particular community's isolation may be a compound of several factors, such as economic practices as well as access to the media or technology. Or, a group of speakers' psychological isolation might be dependent on the extent of its cultural isolation. Take, for example, the Lumbee Indians of Robeson County, North Carolina, who are discussed in more detail below.

This community of speakers has struggled to maintain a distinctive American Indian culture since the 1835 passage of the Revised State Constitution of North Carolina, which mandated that people of color did not have the rights and privileges afforded those who were white. This legislation effectively deprived the Lumbee of their ethnic identity as American Indians since the Lumbee no longer had a discrete identity in the eyes of the government; instead, they were grouped with other people of color.

Such a classification has had an essential effect on how they have defined themselves in relation to other groups. In an important sense, the Lumbee historically have been caught between whites and blacks, and such an indeterminate status has served to foster a sense of otherness apart from both white and black.

In addition to physical or geographic remoteness, sociological or dense, multiplex networks, and economic autonomy, supplemental classifications of isolation, such as community size and historical continuity have also been offered. Isolated communities, indeed, tend to be small in population size (e.g., a few hundred); however, this is not a categorical qualification. For example, the insular community of Lumbee Indians of Robeson County consists of upwards of 45 000 members.

Further, even though many isolated communities may have had significant out-migration and some in-migration over the course of their histories, the historical continuity of the communities tend to remain constant at their core. Most isolated communities have a central core of members who are able to trace back their family lineage to the local inception of the community. More than that, however, true membership of these isolated communities is often tied to such lineage (i.e., whether a speaker's parents, grandparents, and great-grandparents were raised in the area).

While general categories for isolation are critical in the classifications of insular language varieties, the picture is even more complex. With respect to sociological isolation, for example, it is equally important to investigate the various types of contact that are involved as it is to identify the levels of contact facing an insular community (Schilling-Estes, 2002). So, for example, is the contact related to personal associations (e.g., friends, social clubs) or business or economic factors (i.e., work)? Is that contact passive or active (e.g., one-on-one communication or television); mandatory (e.g., one is required by commitment) or optional (e.g., because of personal choice)? Additionally, where does that contact take place?

Is the contact within or outside of the particular community in question?

Types of Isolated Communities

In addition to the classifications that are associated with isolated communities, there are also varying types of isolated communities. That is, open vs. closed and endocentric vs. exocentric communities. Andersen (1988) maintains that it is not unusual for open communities (e.g., those that allow in-migration) to be attitudinally closed, or for closed communities to be open in terms of a wider view of the world. Endocentric communities, however, tend to remain focused on their own internal norms, whereas exocentric communities may be aware of norms outside of their own community reference. Increasing contact with isolated communities, then, may exacerbate or blur lines of distinction depending on psychological factors (Wolfram, in press). As Andersen (1988: 74–75) states, "It may be primarily an attitudinal shift from endocentric to exocentric which changes the course of development of a local dialect when it becomes part of a wider socio-spatial grouping [i.e., when it becomes more open] and not just the opening up of new avenues of interdialectal communication."

Criticism of the Concept of Isolation

Isolation, in whatever way it is categorized, is not a dichotomous classification. In other words, communities are not simply either isolated or nonisolated. Rather, isolation is relative over time, geography, and social space, even in the most extreme dialectical situations.

However, criticism has been leveled at linguists both from within the linguistics discipline and from outside (e.g., historians of Appalachia – see Montgomery, 2000) pertaining to the validity of isolation as a descriptive classification or as an explanatory tool. Many sociolinguists feel that adequately describing the sociohistorical and sociocultural context of a particular community is enough to satisfy an understood set of qualities related to isolation. However, such *ipso facto* classifications can miss the crucial differences that classify very different sociocultural isolated communities. Take, for example, Tristan da Cunha (Shreier, 2003). The island of Tristan da Cunha is located along the maritime mountain ridge, which divides the Atlantic Ocean longitudinally. The closest settlement to the island is Jamestown, St. Helena, about 1400 miles away. There have been periods in this island's history where there was no mail service at all for at least a decade, and today, while they have more contact with the outside world than they ever had (e.g., satellite telephone, television, Internet), this contact is relatively new – since the mid-1990s.

Such isolation is self-evident for this community and has unique linguistic consequences. For example, *useta* + past tense verb (*I useta came*), is a structure distinctive to this community of speakers, created in the learning process of English in a multicontact community. However, this feature has become a dynamic, surviving construction precisely because of the island's fertile isolation.

Similarly, it has been suggested that the extraordinarily high density of discrete languages in Papua New Guinea is at least partly due to the very rugged terrain and historical isolation of speech communities there. For other communities, such as Appalachia, where mountainous regions and pockets of remoteness less clearly separate its speakers from surrounding language varieties, though, such descriptors of isolation may be deceptive and assume a solidarity of the region and stagnancy of the language that does not exist.

Those in favor of rigorously re-investigating the properties of isolated communities assert that sociolinguists have fallen into a trap of treating all isolated communities as the same without regard to the situated context of those communities within the wider scope of language contact. In other words, linguists have not typically asked the question of isolation as compared to what? Why should one community be more isolated than another? Furthermore, under close examination, are all those communities labeled as isolated truly that isolated after all?

Despite critical investigation into the premise and usefulness of isolation as a linguistic construct, however, the role of isolation (or conversely mobility) in dialect development, maintenance, and change has been continually reinforced in a wide array of research.

The Role of Isolation in Language Change

Two principles have historically worked together to situate and describe isolated dialect situations. They are the founder principle and the relic assumption. The founder principle (Mufwene, 1996, 2001) maintains that the linguistic characteristics of a particular speech community are determined by the language that was originally brought into the region. Thus, Appalachian English, for example, is strongly influenced by the language(s) of its original inhabitants (e.g., the Scots-Irish). Likewise, in the case of African American Vernacular English, the issue of donor dialects is still under considerable investigation. The relic

assumption (Andersen, 1988) presumes that isolated dialects will be conservative in their trajectory of linguistic change and will actually remain relatively unchanged after they are clearly formed.

Although there is undeniable and copious support for both of these principles, emerging evidence of isolated communities dampens the principles' associated underlying assumptions:

1. that isolated communities are stagnant with respect to change;
2. that as isolated communities come into contact with surrounding varieties, those communities' distinctiveness disappears;
3. that isolated communities progress on the same trajectory of language change in the midst of outside contact.

The Study of Lumbee English Perfective *Be*

The example of Lumbee English perfective *be* – that is, forms of *be* that are used in aspectual contexts where current varieties of English typically use the perfective form *have*, as in sentences such as *I'm been there before* for *I've been there before* or *You're got my number* for *You've got my number* – yields strong contradiction to the assumptions that isolated communities are stagnant and that they lose their distinctiveness as language contact increases.

The Lumbee Indian Community of Robeson County, North Carolina, have lived in a tri-ethnic situation for close to two centuries with African and Anglo-Americans in Robeson County, NC, located in the southeast corner of North Carolina along Interstate 95 by the South Carolina border.

The county's ethnic populations are near equal proportions, with the Lumbee comprising the largest population of upwards of 45,000. Although no one knows the exact relationship between these three ethnic groups in the earliest contact of the county area, we do know that for well over a century, the ethnic groups have remained separate and distinct. *De facto* segregation exists even beyond the most recent effort to integrate the groups, which involved the desegregation of the county schools in the early 1970s.

The ancestral languages(s) of the Lumbee – as with many American Indian peoples – is unrecoverable: there are no vestiges of American Indian source language features in their current variety of English. Sociohistorical and archaeological evidence does, however, suggest that northern Iroquoian (primarily Tuscarora), Siouan, and Algonquian languages may have had a formative influence. Indeed, it is quite possible that the Lumbee are descendants of a multitribal and/or multiethnic contact situation rather than members of one exclusive native tribe.

Reportedly, the Lumbee were speaking English in their first contact with European immigrants who began inhabiting the Robeson County area around 1730. Where the Lumbee learned their English is still a matter of speculation. Highland Scots (who migrated along the Cape Fear River and spoke varieties of Scottish Gaelic and English) and Ulster Scots (who filtered down into the Piedmont of North Carolina from Pennsylvania and Virginia and spoke a Scots-Irish variety of English) were the principal European language groups of the Robeson County area. However, African-Americans – reportedly brought into the area as slaves by the European groups – were also a significant group in the county.

The current language profile of the Lumbee demonstrates selective alignment of language features with regional and ethnic contact varieties of English, such as general Southern, Appalachian English, and African American Vernacular English. At the same time, however, the Lumbee maintain a few relic features that reflect earlier versions of their English contact varieties. One such feature is perfective *be*.

Perfect *be* is exclusively used by the Lumbee in the context of Robeson County, making the form a localized marker of ethnicity for this group. The use of perfect *be* is widespread in the Lumbee community, spanning all age groups, including both the perfect uses of *is* and *are* (Dannenberg, 2003).

The perfect uses of *be* do not appear to be lexically restricted to co-occur with specific verbs or restricted to particular phonological contexts. Syntactically, perfect *be* may appear in an assortment of contexts, including in embedded clauses (e.g., *I don't know whether I'm still got it or no* or *When you got people, you're got resources*), which demonstrates its flexibility. At the same time, however, perfect *be* is syntactically obligated to occur immediately after the subject in its conjugated forms, and that subject tends to be the first person singular pronoun *I*. Although perfect *be* does, albeit rarely, occur with noun phrases as with other pronominal subjects, such as *we* and plural *you*.

Perfect *be* is a robust form in Lumbee English. Thus, it is not surprising that there are extended and analogous uses of *be* in this variety. For example, there are instances where the perfect *be* actually marks as simple past construction, as in *That's what we're called it, Baby Ruth* – i.e., *That is what we called it when we were little*. In addition to the simple past context, there are also cases where the nonconjugated form of perfect *be* is used in perfect and

simple past constructions, as in *He don't know I be went up to the courthouse and took me some papers out on him* and *I told her don't get it cause Mama be got it*, respectively. Lastly, there are cases of perfect *be* where a hypothetical and/or counterfactual context is marked. For example, when the construction, *They be take their honeymoon there* was uttered, the couple had not married yet, nor had they decided a honeymoon destination. In another example, *We bout be bought all of them for $50 dollars*, the buyers had not actually purchased the items for the sum listed.

Perfect *be* for Lumbee English is a relic form, a retention of an earlier form offered by early contact varieties of English, but relic, here, does not connote static. Indeed, the case of perfect *be* reflects the complexity of a language system as it has configured and reconfigured itself over time.

The Study of Smith Island, Maryland

As the case of Lumbee use of perfective *be* disputes the notion that isolated communities are stagnant with respect to language change, so does the case of Smith Island. Smith Island is located off the coast of Maryland, approximately 10 miles from the mainland Delmarva Peninsula. Schilling-Estes (2002) reports that the island is only accessible by day via a private ferry service, and that the community is largely homogenous. The historical continuity of the residents of Smith Island is also indisputable, with many residents able to trace back their family's lineage to the 1600s. Moreover, Smith Island residents have been able to subsist autonomously in maritime-related activities, although that is changing at present. At the same time that residents are being forced to seek jobs outside of the community (and off the island), the islanders "continue to remain highly focused socially and attitudinally on their local ties" (Schilling-Estes, 2002: 66). Islanders, who sought a high-school education, at one time stayed on the mainland during the school week, as there was no high school on the island. Today, the children are being shuttled to and from the mainland for school on a daily basis.

Smith Island is by all classifications an isolated community, even though current islanders are coming into more and more contact with mainlanders, and in their 1999 study, Wolfram and Schilling-Estes examined two variables in order to demonstrate the level of change in the community over time. The two variables in question were raised /ay/ (as in r*oi*ght for r*i*ght) and glide-fronted /aw/ (as in s*ai*nd for s*ou*nd) (Wolfram *et al*., 1999).

Investigating generational uses (those born between 1899 and 1932, those born between 1942 and 1961, and those born between 1965 and 1971) of both variables from interviews conducted in 1983, Schilling-Estes and Wolfram found that both the raised /ay/ and the glide-fronted /aw/ had increased quite significantly over the three generations. Crucially, they found that these variables actually increased when used in the younger generations; in fact, older speakers had low levels of these features. Thus they concluded that raised /ay/ and glide-fronted /aw/ were not relic features, retained from earlier vestiges of the language, but rather innovative uses that demonstrated dialect concentration rather than dissipation as this community came into more contact with surrounding mainland varieties.

This pattern of concentration holds true for Smith Island with respect to morphosyntactic variables as well. An example is the past *be* paradigm (that is, for negative contexts, leveling to *wasn't* or *weren't* as in *I wasn't/weren't there* and for positive contexts, leveling to *was*), examined in Schilling-Estes (2002). The data illustrate that there was a complete shift of *wasn't* and *was* leveling from the earliest generation studied, where the be paradigm use reflected the typical pattern of usage in vernacular American varieties, to the third generation, where there was categorical *weren't* use for negative contexts (that is, 100%) and some regularization to *was* in positive contexts (22%). As Schilling-Estes (2002) reports, gender differences in the uses of past *be* exacerbate the differences between the generations. In fact, although it took women three generations to switch to categorical usage of *weren't* in negative contexts, the men reached the shift in only two generations. This rapid shift from one usage to the next for male Smith Islanders may be tied to outside contact: men had wider networks than women on the island.

Thus, with the case study of Smith Island, we have further seen that isolated communities can indeed be quite innovative and dynamic with respect to language change. Moreover, in the case of the Lumbee and Smith Island, we have also seen that communities can remain distinctive, even in the face of language encroachment. The Lumbee have selectively carved out a distinctive niche in the context of a tri-ethnic environment and successfully negotiated their identity over time and social space with respect to perfect *be*. Smith Islanders have not only remained unique in the midst of increasing contact with outside communities, but have actually increased their distinctiveness through the concentration of phonological and morphological features. But what do we make of the other assumption that isolated language communities progress on the same trajectory of language change in the midst of outside contact? As the case studies of African American Vernacular English in

Beech Bottom and Hyde County, North Carolina, directions of language change in isolated communities can be quite disparate. Focusing on these two separate case studies in North Carolina highlights that even in a specific geographic region, trajectories of language change may be quite different.

The Case of Hyde County, North Carolina

Hyde County, North Carolina, is located on the eastern seaboard of North Carolina and encompasses coastal mainland as well as the Outer Banks barrier islands. Mainland Hyde County, the geographical region of the study is a long-term biracial community that has been historically one of the lowest populated counties in the state. Currently, there are more than 5000 residents in the county. The terrain of the county is marshy wetlands and the county has been socially detached – in fact, many African adults interviewed reported that they had not left the county until adulthood. In their 2002 investigation into the properties of African American Vernacular English (AAVE) use in Hyde County, Wolfram and Thomas (2002) examined several features well-documented in other studies of AAVE as well as European contact varieties in the Pamlico dialect area in order to situate Hyde County AAVE within regional and wider contact norms. They examined r-lessness (e.g., *cah* for *car*), several phonological variants – including /ai/ (*tide*), /au/ (*house*), /o/ (*boat*), /u/ (*boot*), /e/ (*day*), /ɔ/ (*bought*) – third person s-marking (*The dogs barks*), copula absence (*She going nuts* or *They going nuts*), and past tense *be* regularization (negative *weren't* and positive *was*).

Analysis of each of these features shows a mixed alignment. Clearly, older African-Americans within the county tend to align with their European American cohorts in the county. Older African-American speakers' rates of postvocalic r-lessness and the vowel systems of the speakers solidify this alignment. At the same time, though, younger speakers within the county appear to be moving away from the more traditional vowel systems and aligning with the wider AAVE norm, even though they still show vestiges of the local norm. For example, r-lessness is on the increase from the older generations to the younger generations, but the youngest group still has lower frequency of r-lessness as compared to general AAVE. Moreover, fronted /o/ among younger African-American Hyde County speakers continues to mark them as separated from other AAVE communities.

With respect to the morphosyntactic variables, the alignment is less clearcut. The use of verbal -s, where typically third person -s attaches to plural subjects that tend to be noun phrases as opposed to pronominals, indicates that elderly African-American speakers preferred verbal -s attachment, regardless of subject type, while Hyde County European speakers followed the general pattern, selecting verbal -s with noun phrases. The absence of verbal -s on third person verbs is a well-established descriptive feature in general AAVE. The contrast between European and African-American speakers with respect to this feature is quite clear. All generations of African-American speakers in Hyde County regularly delete third person -s where European American speakers do not.

Copula absence delineates a similar pattern. Among African-American speakers, we would expect a high rate of the copula with *is* as well as a high rate with *are* as per the typical pattern of copula absence in AAVE. Further, we would expect that European American speakers would have some *are* absence, while we would not expect *is* absence. Wolfram and Thomas (2002) report that African-American copula absence in Hyde County is distinctive to that group. Indeed, European Americans in the county do not even have copula absence to a limited degree.

Leveling to *was* and *weren't* in Hyde County patterns are much the same as the above phonological variables. Older African-Americans level to *was* in positive contexts and *weren't* in negative contexts, mirroring European contact variety norms. However, young African-American speakers level to *was* in both positive and negative contexts, much more in line with current varieties of AAVE.

The case study of Hyde County illustrates that in this isolated community, younger African-American speakers are aligning with an outside norm, diverging from their European American cohorts in the region. This trajectory of change is quite different from another African-American community of Beech Bottom, North Carolina.

The Case of Beech Bottom, North Carolina

Beech Bottom is a very small community of self-reported multiethnic speakers (speakers claim European and mixed African-American, European American, and American Indian status) in the southern Appalachian mountain range in Avery County. The community is geographically detached from larger towns and cities and the mountainous terrain hinders access to locations outside of the immediate region. The community is closely knit, and there are only a half-dozen long-term residents of the area. The isolation of this community is certainly without question.

Like the Hyde County study, the investigation of Beech Bottom examines several variables in order to situate speakers within the wider context of local and regional contact norms. Among those features studied were 3rd person plural -s absence and attachment, copula absence, copula absence and past tense *be* regularization, consonant cluster reduction (as in *wes*' for *west*), r-lessness, /ai/ glide reduction (as in *tahm* for *time*), and an overall analysis of the vowel system of representative speakers.

Overall with respect to -s attachment, (Mallinson and Wolfram, 2002) reports that African-Americans and European Americans basically are aligned. Although there are subtle differences in some constraint orders, it appears that -s attachment is declining for both groups over time. At the same time -s attachment demonstrates alignment, 3rd person -s absence illustrates ethnic disparity among the older generations: older European American speakers have a low rate of absence while older African-American speakers have much higher rates of absence. Younger generations of African-American speakers, however, appear to have much lower rates of absence, suggesting a movement toward the local European American variety.

Past *be* also indicates alignment between the groups, with both leveling to *was* in both positive and negative contexts, which is typical of the Appalachian region. Further, regarding r-lessness, the groups are quite similar – that is, both have a high incidence of rhoticity, and both groups reduce the glide on /ai/, both of which are also characteristic of Appalachian varieties.

Conversely, copula absence illustrates dissonance between the groups: African-Americans show much higher rates of absence that their European cohorts. The same holds true with consonant cluster reduction where there is a strong ethnolinguistic divide between the European American and the African-American speakers.

Thus, in this isolated, dwindling community of speakers, African-American speakers are moving toward local European American norms. This movement is clearly contradictory to the Hyde County example, where African-American speakers are on a path of change toward wider AAVE norms. In two isolated communities, then, there are very disparate trajectories of change in the face of contact and change.

Accounting for Relative Variation of Isolated Communities

With obvious differences in the patterns of language variation, maintenance, and change from one isolated community to another, are there common properties that sociolinguists can draw on with respect to concepts of isolation and its effect on language change? Toward an understanding of the complexities and relativity of the concept of isolation, Wolfram (in press) outlines seven principles that may apply to isolated dialect communities:

1. The Principle of dialect exclusion. This principle is designed to account for those dialects that are separated from continuing dialect diffusion among socially dominant groups. Through exclusion, such isolated dialects may have selective retention of features and/or independent language change.
2. Principle of selective change. Under this principle, dialects may exploit dialect structures that may result in divergence. These dialects may be conservative with respect to change or the changes could be accelerated.
3. Principle of regionalization. While independent language change takes place in peripheral dialect communities, isolated dialects may converge with or diverge from other isolated dialects as well as from core dialect communities. This process leads to regionalization.
4. Principle of social marginalization. This principle states that when nonmainstream dialects are marginalized, then those dialects occupy a low status in comparison to socially ratified dialects. Thus, the features of those dialects will be socially stigmatized.
5. Principle of vernacular congruity. There are natural linguistic properties, such as analogical leveling, regularization, and generalization that facilitate language change. Given that those communities that are socially disenfranchised will presumably have less pressure to preserve irregular structures that are socially preferred, such change may unite otherwise disparate communities.
6. Principle of peripheral community heterogeneity. This principle suggests that small communities are quite heterogeneous with respect to intracommunity variation. Such variation may be correlative of the specified ethnic or social boundaries of the community, or the variation could be a result of systematic individual variation. Regardless, small communities can tolerate substantial, long-term variation.
7. Principle of localized identity. Above all, speakers tend to embrace distinctive symbolic tokens of local identity, resulting in the perpetuations of dialect features beyond the enclave community. This perpetuation may be in the form of dialect focusing or intensification.

Isolated language varieties are complex entities of sociolinguistic study. How to precisely identify, categorize, and systematize these kinds of communities of study is still a matter of considerable debate. Out of this debate, however, have come several clear classifications of what isolated dialects actually are and formative principles by which to measure the efficacy of isolated communities as areas of study. Despite controversy along the way, isolated language varieties are the window into lingering questions pertaining to language actuation, embedding, maintenance, and change.

See also: Copula Variation; Cultural and Social Dimension of Spoken Discourse; Endangered Languages; Ethnolinguistic Vitality; Identity and Language; Language/Dialect Contact; Native Speaker; Reversing Language Shift; Social Psychology and Language; Spatial Variation (Geolinguistics); Variation in Native Languages of North America.

Bibliography

Andersen H (1988). 'Center and periphery: adoption, diffusion, and spread.' In Fisiak J (ed.) *Historical dialectology.* Berlin: Mouton de Gruyter. 39–83.

Chambers J (1995). *Sociolinguistic theory.* Cambridge: Blackwell.

Chambers J (2002). 'Dynamics of dialect covergence.' *Journal of Sociolinguistics 6(1),* 117–130.

Childs B (2000). 'The role of contact in isolated transplant communities: the case of consonant cluster reduction in the Bahamas.' Presentation at Southeastern Conference on Linguistics, Oxford, Mississippi, April 2000.

Dannenberg C (2003). Publication of the American Dialect Society 87. *Sociolinguistic constructs of ethnic identity: the syntactic delineation of a Native American English variety.* Durham, NC: Duke University Press.

Dorian N C (1981). *Language death.* Philadelphia: University of Pennsylvania Press.

Dorian, N C (1994). 'Varieties of variation in a very small place: social.

Gauchat L (1905). 'L'unité phonétique dans le patois d'une commune.' In *Aus Romanischen Sparchen und Literaturen: Gestschrift fur Heinrich Morf, Halle: Max Niemeyer.* 175–232.

Grenoble L A & Whaley L J (eds.) (1998). *Endangered languages: current issues and future prospects.* New York: Cambridge University Press.

Mallinson C & Wolfram W (2002). 'Dialect accommodation in a bi-ethnic enclave community.' *Language and Society 31(5).*

Montgomery M (1989). 'Exploring the roots of Appalachian English.' *English World-Wide 10,* 227–278.

Montgomery M (1996). 'Was colonial American English a koine? Studies in English dialectology in memory of Ossi Halainen.' In Klemola J, Kyto M & Rissanen M (eds.) *Bamberger Betrage zur Englische Sprachwissenschaft 38.* Frankfurt: Peter Lang. 231–234.

Montgomery M (2000). 'Isolation as a linguistic construct.' *Southern Journal of Linguistics 24,* 41–54.

Mufwene S (1996). 'The founder principle in creole genesis.' *Diachronica 12,* 83–134.

Mufwene S (2001). *The ecology of language evolution.* Cambridge: Cambridge University Press.

Nagy N (2001). 'Stress and Schwa in Faetar.' In Repetti L (ed.) *Current Issues in Linguistic Theory Series. Italian dialects and phonological theory.* Philadelphia: Benjamins. 239–254.

Poplack S (ed.) (1999). *The English history of African American English.* Malden: Basil Blackwell.

Poplack S & Tagliamonte S (2000). *African American English in the Diaspora.* Malden: Blackwell.

Schilling-Estes N (2002). 'On the nature of isolate and post-isolated dialects: innovation, variation and differentiation.' *Journal of Sociolinguistics 6(1),* 64–85.

Schilling-Estes N & Wolfram W (1999). 'Alternative models of dialect death: dissipation vs. concentration.' *Language 75,* 486–521.

Schrier D (2003a). 'Insularity and linguistic endemicity.' *Journal of English Linguistics 31(3),* 249–272.

Schrier D (2003b). *Isolation and language change: sociohistorical and contemporary evidence from Tristan da Cunha English.* Basingstoke: Palgrave/Macmillan.

Tagliamonte S (1997). 'Obsolescence in the English perfect? Evidence from Samana English.' *American Speech 72,* 33–68.

Thomason S G (2001). *Language contact: an introduction.* Washington, DC: Georgetown University Press.

Trudgill P (1999). 'New-dialect formation and dedialectalization: embryonic and vestigial variants.' *Journal of English Linguistics 27,* 319–327.

Wolfram W (1996). 'Delineation and description in dialectology: the case of perfective I'm in Lumbee English.' *American Speech 71, 5*–26.

Wolfram W (in press). 'The sociolinguistic construction of remnant dialects.' In Fought C *et al*. (eds.) *Sociolinguistic variation.*

Wolfram W & Thomas E R (2002). *The development of African American English.* Malden: Basil Blackwell.

Wolfram W & Schilling-Estes N (2003). 'Remnant dialects in the coastal United States.' In Hickey R (ed.) *The Legacy of Colonial English: a study of transported dialects.* Cambridge: Cambridge University Press.

Wolfram W, Hazen K & Schilling-Estes N (1999). Publications of the American Dialect Society 81. In *Dialect maintenance and change on the Outer Banks.* Tuscaloosa: University of Alabama Press.

Israel: Language Situation

M Amara, Bar-Ilan University, Ramat-Gan, Israel

In 2003, the population of Israel was 6 748 400 (77% Jewish, 19% Arabs, and 4% other minorities). About half of the Jewish people are Ashkenazi and half are Sephardi. The literacy rate is 95% (population percentage over age 15 with ability to read and write). Israel is a highly multilingual country. The *Ethnologue: languages of the world* (see Grimes, 2000) lists 36 languages for Israel; of these, 33 are living languages and 3 are extinct. Israel is a country based on immigration, which partly explains its language diversity.

Hebrew and Arabic are the two official languages of Israel. It is estimated that 4 500 000 Israelis have functional competence in Hebrew, as opposed 2 000 000 who are competent in Arabic (Spolsky and Shohamy, 1999). Hebrew is the dominant language in all walks of life. It is taught as a mother tongue in most of the Jewish schools (except in some ultra-orthodox schools, where Yiddish is the language of instruction) in all stages of education. Hebrew is also learned by all of the minorities as a second language from the 3rd grade on. Modern Hebrew is the national language of the Jews.

Israel is a country based on immigration, therefore the dialects in Israel depend on the origin of parents, and not on geographical location. Israelis who came from Oriental countries before or during the 1950s spoke a distinctly Sephardic (Oriental) variety of Hebrew. However, most second- or third-generation Israelis from these origins shifted to the Ashkenazi Hebrew, so that the Oriental dialect is disappearing except as a socially marked variety.

Arabic is the mother tongue and the main national language of more than 1 million Arab citizens of Israel. Arabic is likewise a community language of hundreds of thousands of Sephardic Jews. Arabic in Israel is a minority language and is legally recognized as a second official language. Arabic is taught as mother tongue in all Arab schools in Israel, from the 1st to the 12th grades, and in several institutes for teacher training. It is also learned by hundreds of thousands of Jews as a foreign language (Amara, 2002; Amara and Mar'i, 2002).

Though Hebrew and Arabic are recognized official languages in Israel, the state of Israel and its society are not publicly bilingual. The status accorded to Arabic in Israeli law is still devoid of any practical significance in Israeli public life. For all intents and purposes, Hebrew is virtually the only language of Israeli civic life. By and large, this is the language in which the Israeli public domain 'speaks'. It is the language of bureaucracy, of higher education, and almost exclusively of the domestic electronic media, and, most importantly, it is the language of those sectors of the labor market that are open to the minority population. The main significance of the status of Arabic in Israel appears, then, not in regard to the society as a whole, but to be the extent of the protection it affords to the internal life of the minority population, especially in regard to the right to education in the minority tongue (Saban and Amara, 2002).

English is the most commonly used foreign language in Israel. It is spoken in both public and private domains (e.g., higher education, electronic media, street signs, and shop signs). The country's laws are also published in English translation, with some delay. English is a compulsory subject from the 3rd or 4th grade on. English is the mother tongue of more than 100 000 Israelis. French is taught optionally as a second foreign language from 5th to 12th grades. Because of the special status of French in the world for political, commercial, social, and cultural reasons, French may be chosen as the second language in place of Arabic.

Immigration remains a major influence on the Israeli language landscape. Recent immigrants, mainly the million Russians and 75 000 Ethiopians, still use their immigrant languages in the home and in the immediate neighborhood. Russian is offered as an optional language for new immigrants through the system. Many established immigrants maintain their own languages (e.g., Polish, Romanian, French, German, Arabic, or Turkish) in their homes and in community settings.

The many Jewish languages that flourished in the Diaspora (such as Ladino, Yiddish, and Judeo-Arabic, with many varieties) are losing ground to the dominant language of the country. Except for Yiddish, maintained in some ultra-orthodox communities, Jewish languages are disappearing in Israel.

In addition to Arabic, which is the major minority language, several minority languages used by non-Jews are reported in Israel. The prominent ones are Armenian, Domari (Nawari), and Adyghe.

Non-Israelis, tourists, and foreign workers (e.g., Thai, Chinese, Filipino, and Romanian) use their own languages, and when they cannot use them or Hebrew, they try to use English as a substitute. Pidginized Hebrew is reported among foreign workers (Elhanan-Granik, 2001).

See also: Arabic; Hebrew, Israeli; Immigrant Language Education.

Bibliography

Amara M (2002). 'The place of Arabic in Israel.' *International Journal of the Sociology of Language 158,* 53–68.

Amara M & Mar'i A (2002). *Language education policy: the Arab minority in Israel.* Dordrecht, Netherlands: Kluwer Academic Publishing.

Elhanan-Granik A (2001). Israeli foreigner talk: employers' verbal communication with guest workers from Romania and Thailand. MA thesis, Bar-Ilan University.

Grimes B F (ed.) (2000). *Ethnologue: languages of the world* (14th edn.) (vol. 1). Dallas, TX: Summer Institute of Linguistics.

Saban I & Amara M (2002). 'The status of Arabic in Israel: reflections on the power of law to produce social change.' *Israel Law Review 36(2),* 5–39.

Spolsky B & Shohamy E (1999). *The languages of Israel: policy, ideology and practice.* Clevedon: Multilingual Matters.

Relevant Websites

http://www.cbs.gov.il – The Central Bureau of Statistics of Israel.

http://www.ethnologue.com – Hardcover, CD, and Web versions of languages of the world available.

Italian

A L Lepschy and G Lepschy, University College London, London, UK

Italian is the official language of the Republic of Italy. It is spoken in Italy (including the Republic of San Marino and the Vatican City), and outside Italy, in the Canton Ticino (it is one of the official languages of the Swiss Confederation) and, with different degrees of vitality, in areas such as those of Nizza, the Principality of Monaco, Corsica, Istrian and Dalmatian towns, and Malta. It also survives in Eritrea, and where there are large communities of Italian emigrants, particularly in the United States (about four million speakers), and in Canada, Argentina, Brazil, and Australia (about half a million each).

Number of Speakers

The inhabitants of the Italian Republic numbered about 57 million in 2002. The number of speakers of Italian is, however, more difficult to establish, mainly because there is no simple way of relating the use of the Italian language with Italian ethnic origin or cultural allegiance, in the case of emigrants, and even with Italian citizenship in the case of the residents of Italy. For the latter, the problem is posed not so much by the linguistic minorities within the boundaries of the Italian republic (speakers of German, ca. 280 000, mainly in South Tyrol; of Occitan and Franco–Provençal, ca. 115 000; of Slovene, ca. 53 000; of Serbo–Croatian, ca. 3000; of Albanian, ca. 100 000; of Greek, ca. 30 000; of Catalan, ca. 15 000), but rather by the presence of the so-called 'Italian dialects.'

Genetic Relationships

The term 'dialect,' in the Italian tradition, is used to refer not to different varieties of the same language but, rather, to 'siblings' of Italian. Italian is based on the literary Tuscan (more specifically Florentine) of the fourteenth century. This in turn derives from Spoken Latin and is therefore a Romance language. But Latin during the period from the sixth to the ninth century gave origin, in Italy, to a myriad of Romance languages, which can be classified into over 15 major groups (broadly coinciding with the Italian regions) such as Piedmontese, Lombard, Venetian, Ligurian, Emilian, and so on. These Italian 'vernaculars', in Italian 'volgari' (a term used to designate the living languages, in the Middle Ages, as opposed to Latin, which was no longer native to any group of speakers, but was the standard written language) in their modern forms constitute the Italian dialects (Sardinian and Friulian are sometimes classified as Italian dialects, sometimes as separate Romance languages). The differences between these dialects can be very pronounced. Speakers of Bolognese and of Neapolitan find each other's dialects as unintelligible as would, say, speakers of French and Spanish.

History

External History

Literary Florentine, as used by the great Tuscan writers of the fourteenth century, Dante, Petrarch, and

See also: Arabic; Hebrew, Israeli; Immigrant Language Education.

Bibliography

Amara M (2002). 'The place of Arabic in Israel.' *International Journal of the Sociology of Language 158,* 53–68.
Amara M & Mar'i A (2002). *Language education policy: the Arab minority in Israel.* Dordrecht, Netherlands: Kluwer Academic Publishing.
Elhanan-Granik A (2001). Israeli foreigner talk: employers' verbal communication with guest workers from Romania and Thailand. MA thesis, Bar-Ilan University.
Grimes B F (ed.) (2000). *Ethnologue: languages of the world* (14th edn.) (vol. 1). Dallas, TX: Summer Institute of Linguistics.
Saban I & Amara M (2002). 'The status of Arabic in Israel: reflections on the power of law to produce social change.' *Israel Law Review 36(2),* 5–39.
Spolsky B & Shohamy E (1999). *The languages of Israel: policy, ideology and practice.* Clevedon: Multilingual Matters.

Relevant Websites

http://www.cbs.gov.il – The Central Bureau of Statistics of Israel.
http://www.ethnologue.com – Hardcover, CD, and Web versions of languages of the world available.

Italian

A L Lepschy and G Lepschy, University College London, London, UK

Italian is the official language of the Republic of Italy. It is spoken in Italy (including the Republic of San Marino and the Vatican City), and outside Italy, in the Canton Ticino (it is one of the official languages of the Swiss Confederation) and, with different degrees of vitality, in areas such as those of Nizza, the Principality of Monaco, Corsica, Istrian and Dalmatian towns, and Malta. It also survives in Eritrea, and where there are large communities of Italian emigrants, particularly in the United States (about four million speakers), and in Canada, Argentina, Brazil, and Australia (about half a million each).

Number of Speakers

The inhabitants of the Italian Republic numbered about 57 million in 2002. The number of speakers of Italian is, however, more difficult to establish, mainly because there is no simple way of relating the use of the Italian language with Italian ethnic origin or cultural allegiance, in the case of emigrants, and even with Italian citizenship in the case of the residents of Italy. For the latter, the problem is posed not so much by the linguistic minorities within the boundaries of the Italian republic (speakers of German, ca. 280 000, mainly in South Tyrol; of Occitan and Franco–Provençal, ca. 115 000; of Slovene, ca. 53 000; of Serbo–Croatian, ca. 3000; of Albanian, ca. 100 000; of Greek, ca. 30 000; of Catalan, ca. 15 000), but rather by the presence of the so-called 'Italian dialects.'

Genetic Relationships

The term 'dialect,' in the Italian tradition, is used to refer not to different varieties of the same language but, rather, to 'siblings' of Italian. Italian is based on the literary Tuscan (more specifically Florentine) of the fourteenth century. This in turn derives from Spoken Latin and is therefore a Romance language. But Latin during the period from the sixth to the ninth century gave origin, in Italy, to a myriad of Romance languages, which can be classified into over 15 major groups (broadly coinciding with the Italian regions) such as Piedmontese, Lombard, Venetian, Ligurian, Emilian, and so on. These Italian 'vernaculars', in Italian 'volgari' (a term used to designate the living languages, in the Middle Ages, as opposed to Latin, which was no longer native to any group of speakers, but was the standard written language) in their modern forms constitute the Italian dialects (Sardinian and Friulian are sometimes classified as Italian dialects, sometimes as separate Romance languages). The differences between these dialects can be very pronounced. Speakers of Bolognese and of Neapolitan find each other's dialects as unintelligible as would, say, speakers of French and Spanish.

History

External History

Literary Florentine, as used by the great Tuscan writers of the fourteenth century, Dante, Petrarch, and

Boccaccio, was gradually accepted as the national standard language and finally codified as such at the beginning of the sixteenth century. The '*Questione della lingua*' (question of the language) was a central concern for the culture of the time with ideological and political, as well as linguistic and literary, implications. Against the proposers of contemporary Tuscan, and of a supraregional '*lingua cortegiana*' (language of the courtiers), the solution that prevailed was the one defended by the Venetian humanist, Pietro Bembo. The model he upheld was archaic Tuscan, which did not imply subservience to any of the rival modern political powers in Italy and was consistent with the principles of classicism and imitation then prevailing.

This had important consequences for the linguistic history of Italy. On the one hand Italian appears to be exceptionally 'conservative': it must be almost unique, among modern national languages, in having changed so little through eight centuries of documented history, and in allowing a twentieth-century reader to approach thirteenth-century Tuscan texts with minimum effort. On the other hand (and this is perhaps the condition for such stability) this language remained a standard written model, accessible only through literacy, like Latin, rather than acquired as a native language. The vast majority of the people, who were illiterate, did not have access to the literary language, and in speech used exclusively their native dialect. It has been calculated that at the moment of political unification in 1860 the number of the inhabitants of the country who could use Italian was at most about 10%, perhaps as low as 2.5%. In 2005, after 145 years of existence of a unified Italy, Italian was assumed to be known to all Italians, but statistics indicated that almost one-half of the population still preferred to use a dialect rather than the national language.

Internal History

The main points to be mentioned in the change from Latin to Italian include the following.

Phonology The opposition of long and short vowels in Latin was replaced by a distinction determined by stress and syllable structure: long vowels are obligatory in free stressed nonfinal syllables, short vowels in other conditions (i.e., in checked stressed, free stressed final, and all unstressed syllables). In stressed syllables the five Latin vowel qualities became seven, with long *ē* and *ō* giving midhigh [e] and [o], short *Ĕ* and *Ŏ* giving midlow [ɛ] and [ɔ] (these broke into [jɛ] and [wɔ] in free syllables). In unstressed syllables only five vowels are used: [i], [e], [a], [o], [u]. The consonant system undergoes the following main changes: assimilation (e.g., [kt] > [tt], as in *factum* > *fatto*); palatalization and assibilation before front vowels (as in *cenam* [keːnam] > *cena* [tʃena]); *hodie* > *oggi*, *medium* > *mezzo*); sonorization, which applies unsystematically, as in *stratam* > *strada*, but *amatam* > *amata*. Initial *h* and final consonants (apart from nasals and liquids in proclitics: *un*, *per*, *il*, etc.) are dropped.

Morphology The case system of the Latin declension disappears and prepositions are used instead: *di uomo* replaces Latin *hominis*. The neuter gender disappears. The Latin verb pattern is basically preserved, with the introduction of 'analytical' or compound forms for the passive (*è amato* for *amatur*), the future (*amare ho*, whence the new synthetic *amerò*, for *amabo*), the perfect (*ho amato* for *amavi*, although this survives as *amaï*), and so on.

Syntax The freedom of Latin word order is reduced, as syntactic function is signaled by linear position rather than case endings: for 'Paul saw Peter' the only normal and unequivocal structure in Italian is '*Paolo vide Pietro*,' whereas in Latin *Petrum Paulus vidit* would be equally clear in any of the six theoretically possible combinations of these three words.

Written Records

The first dated text in an Italian vernacular is found in the account of a court case in Campania in the year 960 A.D. (the *Placito capuano*). The original record in Latin has been preserved; it includes the statement repeated by some witnesses in the vernacular: *sao ko kelle terre per kelle fini que ki kontene trenta anni le possette parte sancti Benedicti* ('I know that those lands within those boundaries which here are contained, for 30 years the party of St Benedict owned them'). Several, mostly short vernacular texts, are found subsequently until the thirteenth century, when 'Italian' literature proper begins, with religious compositions, the poetry of the Sicilian School, and finally Tuscan poetry.

The vernacular must have been used in speech for a long time before the document of 960. There are several texts that appear to represent an attempt to fix in writing a kind of vernacularized Latin by lowering it toward the spoken language, or raising the latter to conform to the conventions of written Latin. One of the best-known documents of this 'compromise' is a riddle apparently jotted down in Verona, at the beginning of the ninth century on a page of a prayer book; it refers to the act of writing, under the guise of describing the work of a farmer: '*se pareba*

boues alba pratalia araba & albo uersorio teneba & negro semen seminaba' (a plausible rendering, among many which have been proposed: 'he was driving oxen, he was plowing white fields and he was holding a white plow, and he was sowing black seed'); some features appear more Latin than vernacular (e.g., *-b-* for *-v-* in the imperfects, final *-n* in *semen*, etc.); others are more vernacular (e.g., fall of final *-t* in the imperfects, *e* in *negro*, etc.). However, many traits present in dialects are attributed to the substratum of pre-Latin idioms spoken in Italy at the time of Romanization. For instance, a Neapolitan feature such as *-nn-* for *-nd-* belongs to Oscan and is found in the Latin graffiti of Pompeii, preserved under the ashes from the eruption of Vesuvius in 79 A.D. This, however, does not entitle us to suggest that Neapolitan was spoken in the first century A.D.

Writing System

The Italian writing system derives from the Latin one. It uses the Roman alphabet as it was adapted for the vernacular during the Middle Ages. Palatalization created some problems, which are solved in the Italian spelling conventions, by using – in front of *i* and *e* – *c*, and *g* for the palatals [tʃ] and [dʒ], and *ch* and *gh* for the velars [k] and [g]. But the system does not render all the oppositions of Tuscan phonology. In particular there are individual letters that correspond to contrasting sounds: *e* to [e] and [ɛ], *o* to [o] and [ɔ], *s* to [s] and [z], *z* to [ts] and [dz]; the last two alveolar affricates, which are always long intervocalically, are represented in spelling sometimes by *z* and sometimes by *zz*. These points leave a trace in the history of the language. As Italian was adopted as a national language in its written rather than spoken form, it became established, in different parts of Italy, with phonological counterparts, for these 'ambiguous' letters, which may be different from the Tuscan ones. Considering that contemporary Tuscan does not constitute an undisputed standard, Tuscan pronunciations like *b*[ɛ]*ne*, *ca*[s]*a*, or [ts]*io* are not felt, in present-day Italian, to be more correct than Northern ones like *b*[e]*ne*, *ca*[z]*a*, or [dz]*io*.

Individual Characteristics

Among the Romance languages Italian would appear to be typologically more similar to Spanish than to French.

For syntax, one frequently noted feature is that it is a 'pro-drop' language (i.e., it does not need to express the pronominal subject of a verb and that, consistently, it can put the subject after the verb). The adjective has two positions: postnominal if it is restrictive, and prenominal otherwise.

Morphology is conservative within a traditional Indo–European pattern; adjectives, nouns and verbs are subdivided into classes, without apparent semantic justification. An adjective may behave like *ross-o* (MASC SG), *ross-a* (FEM SG), *ross-i* (MASC PL), *ross-e* (FEM PL); or like *verd-e* (MASC and FEM SG), *verd-i* (MASC and FEM PL); or like *par-i* (MASC and FEM SG and PL). A noun may behave like *la cas-a* (FEM SG) *le cas-e* (FEM PL); or *il poet-a* (MASC SG) *i poet-i* (MASC PL); or *il libr-o* (MASC SG), *i libr-i* (MASC PL), and so on. A verb may behave like *cant-are*, or *ved-ere*, or *dorm-ire*. Each conjugation has a vast array of different forms according to person (*canto*, *canti*, *canta*), number (*canta*, *cantano*), tense (*canto*, *canterò*, and, with aspectual distinctions, *cantava*, *cantò*), and mood (*lui canta*, *che lui canti*). There is a complex pattern of agreement for gender and number that involves nouns, articles, adjectives, and past participles: *è arrivata una ragazza alta, è arrivato un ragazzo alto*. A notable feature in Italian derivational morphology is the productivity of evaluative suffixes (often called 'alteration'): the following forms are based on *libro*: *librone, libretto, librino, libruccio, libraccio*, etc.

Phonology

The rhythm of Italian is syllable timed. There are seven vowels, very similar to the cardinal ones: /i/, /e/, /ɛ/, /a/, /ɔ/, /o/, /u/. In unstressed syllables the opposition of midhigh and midlow vowels is neutralized, the quality of the other vowels remains distinct. There are two semiconsonants: /j/ and /w/, and 21 consonants: /p/, /b/, /t/, /d/, /k/, /g/, /ts/, /dz/, /tʃ/, /dʒ/, /f/, /v/, /s/, /z/, /ʃ/, /m/, /n/, /ɲ/, /l/, /ʎ/, /r/. Typologically uncommon is the systematic opposition of long to short consonants, which applies to all the items listed apart from six: /z/ is always short, and /ts/, /dz/, /ʃ/, /ɲ/, and /ʎ/ are always long intervocalically.

Illustration

Here follows a sentence (1) quoted for illustration:

(1) *A Venezia è più facile che si senta parlare il dialetto che l'italiano.*
'In Venice it is easier to hear people speaking dialect than Italian.'

Phonemic transcription: /a vve'nɛttsja ɛ ppju f'fatʃile ke ssi 'sɛnta par'lare il dja'lɛtto ke ll ita'ljano/. Note that this sentence is pronounced in Venice without 'syntactic doubling' (i.e., the lengthening of initial consonants in specified conditions) and with some of the variations mentioned above in the section,

'Writing System': /a ve'nɛtsja ɛ pju 'fatʃile ke si 'senta par'lare il dja'letto ke l ita'ljano/. In the local dialect this would be: /a ve'nɛsja ze pju 'fasie ke se 'seŋta par'lar el dja'eto ke l ita'ljaŋ/.

Note the gender and number agreement between articles and nouns; the use of the subjunctive governed by *facile che*; the interesting construction with *si*, which can be interpreted as an impersonal (with *si* acting as the indefinite subject of *senta*: 'one hears'), or as a passive, with the infinitive clause (*parlare*) acting as sentential subject ('speaking is heard'). Also, the subject of the infinitive need not be specified: '*si sente parlare*' 'one hears [someone] speak.' There is no dummy subject for *è* ('it is'), and the sentential subject clause introduced by *che* comes after the verb. The first *che* acts as a complementizer, and the second *che* as a conjunction within the comparative structure (*più* ... *che*).

Sociolinguistic Points

As mentioned above, in Italy there are many linguistic enclaves in which 'foreign' languages are spoken, and in some cases their use is 'protected' by special legislation. For the majority of Italians the traditional situation was one of diglossia (with the local dialect used in speech, and literary Italian in writing. After political unification, and particularly as a consequence of far-reaching social changes, such as internal migration (mostly from the south to northern industrial conurbations) and the influence of the mass media, Italian has been widely adopted in speech as well. Regional differentiation is clearly marked in phonology, identifiable in lexis, and less clearly noticeable in grammar. A colloquial variety of the language has developed that has been called 'popular Italian'; this appears to be gaining acceptability, and some of its features are penetrating into the standard written language (e.g., *gli* is now frequently found in writing for 'to them,' and sometimes even for 'to her,' as well as for the traditional 'to him'). The dialects have been diagnosed often as terminally ill and on the point of demise. In fact, they have proved remarkably resilient in ordinary usage, and sometimes they appear to be taken up by people (including the young) as a way of reasserting their own group identity and reacting against an alienating process of national equalization. There has also been a striking vitality in dialect poetry, often using very local, individual forms of the dialect, rather than a generalized, regional variety.

Descendants

Italian played an important part in the formation of the lingua franca used in the past in the Mediterranean. Since the Renaissance there have also been varieties of '*lingua zerga*' employed by vagrants. A level of 'slang' is generally thought not to be popular in Italian because its functions were filled by the dialects. A curious form of English slang, called 'parlyaree' or 'polari,' traditional among sailors, actors, and gays (and now surviving in the form of lexical relics) was based on Italian. Italian-based creoles seem to exist or to have existed in Eritrea, Argentina, and Brazil (Harris and Vincent, 1987: 20).

Languages Influenced

The influence of Italian has been mainly felt, in all the major European languages, at the level of high culture, in the lexis of music and the figurative arts. In the Renaissance, and for a long time subsequently, an acquaintance with Italian was thought to be part of the cultural equipment of educated people in most European countries.

History of Linguistic Investigation

More than for other European languages, linguistic awareness and a discussion of linguistic matters (the *Questione della lingua*) has always been very relevant for Italian intellectuals. During the Renaissance the description and codification of the language was an important part of national culture, and the *Vocabolario* (1612) of the Accademia della Crusca was the first of the great national dictionaries to be published. Less important was Italy's contribution to linguistic studies (including the historical investigation of the Italian language and dialects) during the nineteenth and twentieth centuries.

See also: Indo–Aryan Languages; Indo–European Languages; Italian as a Fusional Language; Italian Lexicography; Italy: Language Situation; Romance Languages.

Bibliography

Cortelazzo M, Marcato C, De Blasi N & Clivio G P (eds.) (2002). *I dialetti italiani. Storia struttura uso*. Turin: UTET.

De Mauro T (1970). *Storia linguistica dell' Italia unita*. Bari: Laterza.

Devoto G (1978). *The Languages of Italy*. Chicago, IL: University of Chicago Press.

Harris M & Vincent N (eds.) (1987). *The Romance languages*. London: Croom Helm.

Holtus G, Metzeltin M & Schmitt C (eds.) (1988). *Lexikon der Romanistischen Linguistik, Bd IV. Italienisch, Korsisch, Sardisch*. Tübingen: Niemeyer.

Lepschy G (2002). *Mother tongues and other reflections on the Italian language*. Toronto: University of Toronto Press.

Lepschy A L & Lepschy G (1991). *The Italian language today.* London: Routledge.
Maiden M (1995). *A linguistic history of Italian.* London: Longman.
Maiden M & Parry M (eds.) (1997). *The dialects of Italy.* London: Routledge.
Migliorini B (1984). *The Italian language.* London: Faber and Faber.

Italian as a Fusional Language

C Iacobini, Università di Salerno, Salerno, Italy

In a fusional language words are inflected (and derived) by using affixes whose boundaries are difficult to identify due to the tendency of affixes to fuse with one another and with the root. The numerous occurrences of allomorphy and the tendency of morphs to simultaneously encode several meanings result in the fact that there is no one-to-one correspondence between morphs and morphemes, thus making linear segmentation difficult.

Other terms are sometimes used instead of *fusional*: *flectional, inflectional*, or *inflecting, inflective*. The term *fusional* should be preferred to others referring to inflection because both agglutinating and polysynthetic languages can be highly inflectional.

Morphological Typology

The fusional type is one of the main morphological types. The classification of languages by morphological types is part of the standard terminology of linguistics, but it is also strongly criticized by the majority of typologists for three main reasons: (1) because the classification criteria are rather vague and difficult to apply in a consistent way; (2) because the morphological type is defined in terms of mutual favorability of properties rather than of implicational correlations, resulting in a low predictive power; and (3) because morphological typology has a holistic background.

The vagueness of the classification in morphological types is shown by the lack of consensus on both the number of types and the number of parameters identifying them (the three most-used parameters are (1) the ratio of morphs to word forms, (2) the number of morphemes to morphs, and (3) the degree of word-internal modification of morphs (Greenberg, 1954).

Modern linguistics has disavowed the ideological prejudice dating back to the 19th century, according to which the fusional morphological type was regarded as being superior to other types. Just like these, the fusional type is a combination of functionally interconnected features, which – as a whole – form an ideal construct characterizing (the whole, or some aspects of) the morphology of languages. Languages are rarely pure types; they usually mix elements of different types. Assigning a language to a specific type depends on the preponderance of features considered significant (the quantification of such features is a difficult problem to solve from a practical point of view).

Despite criticism, the classification by morphological types is convenient and widely used to rapidly identify a number of features that tend to cooccur in the morphology of a language, and also to assess the extent to which a language moves away from such ideal constructs, both in a synchronic and in a diachronic perspective (some authors argue that languages tend to move toward a typological goal (Dressler, 1985).

The Fusional Type

The best-known attempt to establish a list of features that cooccur in morphological types is the one made by Skalička (1966). The features that tend to cluster in languages displaying fusional morphology can be listed as follows:

1. Words are formed by a root and (one or more) inflectional affixes, which are employed as a primary means of indicating the grammatical function of the words in the language. Agreement is widely employed.
2. High degree of modification of internal morpheme boundaries, with a consequently difficult linear segmentation.
3. Tendency to cumulate morphological meanings in a single affix (with consequent asymmetry between the semantic and formal organization of grammatical markers).
4. Word-class distinction is maximal. Inflection is rich, regarding both the number of inflectional classes and the extension of paradigms.
5. Stem suppletion; many cases of both homonymy and synonimy among affixes; clear distinction between inflectional and derivational affixes.
6. A slight correlation with syntax can be seen in the relatively free word order (but there are also fusional languages with a fairly fixed word order).

The fusional type is differentiated from the isolating type by the use of bound morphs and the clear-cut distinction between word classes; it is differentiated

Lepschy A L & Lepschy G (1991). *The Italian language today.* London: Routledge.
Maiden M (1995). *A linguistic history of Italian.* London: Longman.
Maiden M & Parry M (eds.) (1997). *The dialects of Italy.* London: Routledge.
Migliorini B (1984). *The Italian language.* London: Faber and Faber.

Italian as a Fusional Language

C Iacobini, Università di Salerno, Salerno, Italy

In a fusional language words are inflected (and derived) by using affixes whose boundaries are difficult to identify due to the tendency of affixes to fuse with one another and with the root. The numerous occurrences of allomorphy and the tendency of morphs to simultaneously encode several meanings result in the fact that there is no one-to-one correspondence between morphs and morphemes, thus making linear segmentation difficult.

Other terms are sometimes used instead of *fusional*: *flectional, inflectional*, or *inflecting, inflective*. The term *fusional* should be preferred to others referring to inflection because both agglutinating and polysynthetic languages can be highly inflectional.

Morphological Typology

The fusional type is one of the main morphological types. The classification of languages by morphological types is part of the standard terminology of linguistics, but it is also strongly criticized by the majority of typologists for three main reasons: (1) because the classification criteria are rather vague and difficult to apply in a consistent way; (2) because the morphological type is defined in terms of mutual favorability of properties rather than of implicational correlations, resulting in a low predictive power; and (3) because morphological typology has a holistic background.

The vagueness of the classification in morphological types is shown by the lack of consensus on both the number of types and the number of parameters identifying them (the three most-used parameters are (1) the ratio of morphs to word forms, (2) the number of morphemes to morphs, and (3) the degree of word-internal modification of morphs (Greenberg, 1954).

Modern linguistics has disavowed the ideological prejudice dating back to the 19th century, according to which the fusional morphological type was regarded as being superior to other types. Just like these, the fusional type is a combination of functionally interconnected features, which – as a whole – form an ideal construct characterizing (the whole, or some aspects of) the morphology of languages. Languages are rarely pure types; they usually mix elements of different types. Assigning a language to a specific type depends on the preponderance of features considered significant (the quantification of such features is a difficult problem to solve from a practical point of view).

Despite criticism, the classification by morphological types is convenient and widely used to rapidly identify a number of features that tend to cooccur in the morphology of a language, and also to assess the extent to which a language moves away from such ideal constructs, both in a synchronic and in a diachronic perspective (some authors argue that languages tend to move toward a typological goal (Dressler, 1985).

The Fusional Type

The best-known attempt to establish a list of features that cooccur in morphological types is the one made by Skalička (1966). The features that tend to cluster in languages displaying fusional morphology can be listed as follows:

1. Words are formed by a root and (one or more) inflectional affixes, which are employed as a primary means of indicating the grammatical function of the words in the language. Agreement is widely employed.
2. High degree of modification of internal morpheme boundaries, with a consequently difficult linear segmentation.
3. Tendency to cumulate morphological meanings in a single affix (with consequent asymmetry between the semantic and formal organization of grammatical markers).
4. Word-class distinction is maximal. Inflection is rich, regarding both the number of inflectional classes and the extension of paradigms.
5. Stem suppletion; many cases of both homonymy and synonimy among affixes; clear distinction between inflectional and derivational affixes.
6. A slight correlation with syntax can be seen in the relatively free word order (but there are also fusional languages with a fairly fixed word order).

The fusional type is differentiated from the isolating type by the use of bound morphs and the clear-cut distinction between word classes; it is differentiated

from the agglutinating type by the kind of juncture between morphemes and the nonbiunivocal correspondence between morphs and morphemes. In the synthetic vs. analytic distinction, the fusional type tends toward the synthetic end.

The fusional type is largely represented in Indo–European languages, especially the most conservative ones. Latin, Slavonic, and Romance languages are the ones that are most often mentioned as good representatives of this morphological type. The main features identifying the fusional type will be described hereafter using examples drawn from Italian.

Affixal Inflection and Agreement

Italian makes wide use of inflectional affixes (nouns, verbs, adjectives, pronouns, articles are usually inflected, whereas adverbs are invariable). Inflection is obtained by replacing the word ending.

The two following sentences show an example of agreement in the singular and in the plural:

quell-a buon-a ***tort-a*** *è finit-a, l-a avrei mangiat-a volentieri*
'that good pie is finished, I would have eaten it with pleasure'
quell-e buon-e ***tort-e*** *sono finit-e, l-e avrei mangiat-e volentieri.*

All the underlined elements agree in number and gender with the feminine noun in bold, the auxiliary agrees only in number, the adverb is invariable; it should be noted that agreement also affects the pronoun and the participle in the second clause.

Cumulative Exponence in Adjectival Inflection

The main adjectival inflectional class has four endings

bell-o 'nice' (sing MASC)
bell-i (PL MASC)
bell-a (sing FEM)
bell-e (PL FEM)

Each affix codes two grammatical meanings at the same time (i.e., gender and number). Regarding adjectival inflection, the Italian language is more fusional than other Romance languages. In Spanish, for example, the gender and the number are expressed by two separate morphs, and the morph for the plural is the same for the two genders:

hermos-o (sing MASC)
hermos-o-s (PL MASC)
hermos-a (sing FEM)
hermos-a-s (PL FEM)

The other important inflectional class employs two affixes: *-e* for the singular, *-i* for the plural. Homonymy is therefore observed between the affixes from the two classes (*-e* can signify both PL FEM and sing MASC or FEM; *-i* can signify both PL MASC or FEM). Compared with Latin, Italian has lost the ability to form the comparative through affixation (the few affixal comparatives in use are instances of fused exponence: in *migliore* 'better' and *minore* 'minor' it is not possible to segment the lexical base from the comparative suffix), whereas the superlative form is productively obtained by using the suffix *-issimo* (*bello/bellissimo* 'very beautiful').

Noun Inflection and Affixal Homonymy

A recent classification of Italian nominal inflection proposed by D'Achille and Thornton (2003) – superseding the very unsatisfactory traditional classification based on the endings of the singular form of nouns – distinguishes five inflectional classes (defined as a set of lexemes whose members each select the same couple of endings for singular and plural), plus a sixth class consisting of invariable nouns (the bottom line of Table 1 shows the percentage of types of each class in the Italian basic vocabulary).

Class 1 consists of masculine nouns in the overwhelming majority (*libro/-i*), with very few feminine exceptions (*mano/-i*). Class 2 is made up of feminine nouns only. Class 3 consists of masculine nouns by approximately 45% (*fiore/-i*), of feminine nouns by the same percentage (*siepe/-i*), and of ambigeneric by 10% (*cantante/-i* may be used both for a male or a female singer). Class 4 is mostly composed of masculine nouns (*poeta/-i*) but includes a couple of feminine nouns as well (*ala/-i*); in class 5 nouns are masculine in the singular and feminine in the plural (*uovo/-a*); all the nouns ending in a stressed vowel or in a consonant (class 6) are invariable, but class 6 also includes invariable nouns ending in each of the four vowels used as inflectional endings, except for /a/ in feminine nouns (MASC: *cinema, golpe, kiwi, stereo*; FEM: *consolle, ipotesi, radio*). Although there seems to be a tendency to express the singular by

Table 1 Nominal inflectional classes and percentage of types of each class in the Italian basic vocabulary

Nominal Inflectional Class	*1 sing/PL*	*2 sing/PL*	*3 sing/PL*	*4 sing/PL*	*5 sing/PL*	*6 sing/PL*
Endings	-o/-i	-a/-e	-e/-i	-a/-i	-o/-a	invariable
%	41.2%	30.3%	20.6%	1.2%	0.2%	5.4%

means of back vowels (*o* MASC, *a* FEM) and the plural by means of front vowels (*i* MASC, *e* FEM), the cases of affixal homonymy are quite numerous (remember that class 3 uses front vowels both for the singular and for the plural, and it includes masculine and feminine as well as ambigeneric nouns). Although on the one hand, the recent trend toward an increase of invariable nouns ending by a vowel curbs the ratio of inflected nouns, on the other it decreases the degree of correlation between form and inflectional class. The low predictability of both the inflectional class and gender is clearly shown in Table 2, which refers to singular (Thornton, 2001).

Traces of Case in Pronouns

Italian has lost a fusional feature that characterizes Latin nominal inflection: the case. However, one can detect traces of case in the Italian pronominal system. The choice of the correct form of certain pronouns (Table 3) demands a decision that depends on whether pronouns express a subject, a direct, or an indirect complement.

The system of stressed pronouns distinguishes three persons for the singular and three for the plural. The third-person singular has distinct forms on the basis not only of gender, but also of the feature ± human. The third-person plural subject differentiates gender but not the ± human feature. In the spoken language, the pronouns *lui, lei*, and *loro* tend to broaden their functional scope and are frequently used both as subject and nonsubject, even for nonhuman referents.

The Table 4 shows clitic pronouns with the function of direct and indirect complement.

The use of such pronouns is complicated by rules of reciprocal ordering as well as by allomorphy: all clitic pronouns that end in *–i* replace the vowel with /e/ if followed by another clitic.

ti racconto una storia '[I'll] tell you a story'
te la racconto'[I'll] tell it to you'
te lo porto '[I'll] bring it to you'
portaglielo 'bring it to him/her'

The pronoun *le* (singFEM) becomes *gli* when followed by another clitic, which results in homonymy and consequent breakdown of gender distinctions:

mostrale la stanza 'show her the room'
mostragli la stanza 'show him the room'
mostragliela 'show it to her/him'

Homonymy occurs also with other unstressed forms (*ci* and *vi* are also locative adverbs; *lo, la, gli*, and *le* articles) and with stressed pronouns in postverbal position (*te lo prendi con te* 'you take it/him with you').

Verb Inflection

The verb is the word class with the richest inflection both in number of forms and variability (many high-frequency verbs have irregular inflection). Tense, mood, person, and number are expressed affixally.

The most relevant features of the fusional type appear even within regular inflection. Grammatical meanings can be expressed as fused in a single morph (in the form *amo*, 'I love,' there is no other overt expression than the ending-*o* to code indicative mood, present tense, first-person, singular), as well as realized through a combination of several affixes, with complex exponence relations holding between morphs: in *canterébbero* '[if] they would sing' three morphemes (conditional, third person and plural) are signaled by the entire termination-*rébbero* as a whole. A linear segmentation is not possible because -*bb*- occurs only with forms that are both third person and conditional. The final -*ro* expresses plural and again third person, the stressed -*e*- occurs consistently only in conditional forms, and the -*r*- occurs with futures and (again) conditionals (Matthews, 1970).

A suprasegmental modification (stress on the last syllable) suffices to distinguish between two regular verb forms like *tème* (INDIC PRES 3sing) and *temé* (INDIC preterit 3sing).

Table 2 Low predictability of gender distinction on the basis of the singular ending of nouns

-o	prevailing MASC, but also FEM (*mano, foto*)
-a	prevailing FEM, but also MASC (*papa, clima*)
-e	equally divided between MASC (*fiore, cantante*) and FEM (*siepe, cantante*)
other	equally divided between MASC and FEM

Table 3 Stressed personal pronouns

Stressed Pronouns	*1 sing*	*2 sing*	*3 sing*	*1 PL*	*2 PL*	*3 PL*
Subject	io	tu	+hum: egli (MASC) ella (FEM) −hum: esso (MASC) essa (FEM)	noi	voi	essi (MASC) esse (FEM)
Nonsubject	me	te	+hum: lui (MASC) lei (FEM) −hum: esso (MASC) essa (FEM)	noi	voi	loro

Table 4 Clitic personal pronouns

Clitic Pronouns	*1 sing*	*2 sing*	*3 sing*	*1 PL*	*2 PL*	*3 PL*
ACC	mi	ti	lo (MASC) la (FEM)	ci	vi	li (MASC)le (FEM)
DAT	mi	ti	gli (MASC) le (FEM)	ci	vi	-

Table 5 Indicative present regular inflection of the three traditional verbal classes

INDIC PR	*Class-a-sing*	*PL*	*Class-e-sing*	*PL*	*Class-i-sing*	*PL*
1	cànt-o	cànt-iàmo	tèm-o	tem-iàmo	àpr-o	apr-iàmo
2	cànt-i	cànt -à-te	tèm -i	tem-é-te	àpr-i	apr-ì-te
3	cànt-a	cànt -ano	tèm -e	tèm -ono	àpr-e	àpr-ono

Table 5 shows instances of regular inflection of indicative present for the three traditional classes on the basis of the thematic vowel (accent marks have been added, even though not in use in normal orthography). Noteworthy, are the cumulative expression of more meanings by means of a single ending, and the omission of the thematic vowel with the exception of the second plural.

Table 6 Indicative imperfect inflection of regular class -a- verbs

INDIC Imperfect class -a-	*sing*		*PL*
	1	cant-à-v-o	cant-a-v-àmo
	2	cant-à-v-i	cant-a-v-àte
	3	cant-à-v-a	cant-à-v-ano

The thematic vowel may also undergo some allomorphies: in class *-a-* it is replaced by /e/ in the indicative future (*amerò*) and the present conditional (*amerei*), in class *-i-* by /e/ in the present participle (*aprente*) and in the gerund (*aprendo*), in class -e- by /u/ in the past participle (*voluto*), and by /i/ in the derivation of deverbal nouns and adjectives (*temibile*, *spremitura*).

The indicative imperfect is the tense that least identifies with the fusional type, for it shows the best correspondence between morphs and morphemes in Italian conjugation (Table 6). But also in the imperfect the categories of person and number, as well as those of mood and tense, merge in a single morph. The person–number affixes largely overlap (even though not entirely) with those of the present, whereas the morph *-v-* has no other use in the conjugation.

Besides synthetic forms, there are also analytic forms resulting from the combination of an auxiliary and a past participle. The main auxiliary verbs are *essere* 'to be' and *avere* 'to have'; both are highly irregular

INDIC PRES *sono, sei, è, siamo, siete, sono*
INDIC preterit *fui, fosti, fu, fummo, foste, furono*
INDIC IMPERF 1PERS *ero*, INDIC FUT 1PERS *sarò*; PP *stato*

The auxiliary expresses tense, mood, person, and number, whereas the past participle can either agree in gender and number or be employed in the citation form (singMASC), whereas the agreement between participle and subject is systematic with verbs whose auxiliary is *essere* (*i ragazzi sono partiti* 'the boys have left'/*la ragazza è partita* 'the girl has left'), in verbs whose auxiliary is *avere* the agreement is with the object, and it occurs only within limited contexts (*hai comprato le pere? sì le ho comprate* 'have you bought the pears? Yes, I have bought them').

The indicative future and the present conditional are interesting from a typological point of view. They have stemmed from an analytic origin to reach a fusional state.

Many verbs have an irregular conjugation. The majority belong to the *-e-* class, some others to the *-i-* class, only three to the *-a-* class. The traditional classification does not allow to catch the similarities between *-e-* and *-i-* classes. Dressler *et al.* (2003) classify Italian verbs in two macroclasses based on the productivity criterion, as well as on formal correlations that allow us to also take into account the subdivisions within each class.

It would be neither possible nor helpful to mention here all the instances in which the verbs depart from the regular conjugation model. Idiosyncrasies occur primarily in the present (indicative and subjunctive), the preterit, and the past participle.

Pirrelli and Battista (2000) show that in the irregular conjugations the modifications of the stems range along a continuum from minor phonological processes to clear suppletion instances.

Table 7 Examples of stem modifications in the indicative present inflection of some very frequent irregular verbs

	nascere	*'be born'*	*udire*	*'hear'*	*sedere*	*'sit'*	*finire*	*'finish'*	*dovere*	*'must'*	*andare*	*'go'*
	sing	*PL*	*sing*	*PL*	*sing*	*PL*	*sing*	*PL*	*sing*	*PL*	*sing*	*PL*
1	'nasko	na'ʃːamo	'ɔdo	u'djamo	'sjɛdo	se'djamo	fi'nisko	fi'njamo	'devo	do'bːjamo	'vado	an'djamo
2	'naʃːi	na'ʃːete	'ɔdi	u'dite	'sjɛdi	se'dete	fi'niʃːi	fi'nite	'devi	do'vete	'vai	an'date
3	'naʃːe	'naskono	'ɔde	'ɔdono	'sjɛde	'sjɛdono	fi'niʃːe	fi'niskono	'deve	'devono	'va	'vanno

Table 8 Examples of stem modifications in the indicative preterit inflection of some very frequent irregular verbs

	venire 'come'		*perdere 'lose'*		*fondere 'melt'*	
	sing	*PL*	*sing*	*PL*	*sing*	*PL*
1	vénni	venìmmo	pèrsi	perdémmo	fùsi	fondémmo
2	venisti	veniste	perdésti	pérdéste	fondésti	fondéste
3	vénne	vénnero	pèrse	pèrsero	fùse	fùsero

Table 7 displays the phonetic transcription of indicative present conjugation of six irregular verbs that exhibit a number of stem alternations: palatalization before a front vowel (*nascere*), ablauting of root vowel (*udire*), diphthongization (*sedere*), *-isc-* insertion and palatalization (*finire*), ablauting, consonant labialization and lengthening (*dovere*), and stem suppletion (*andare*).

Two things are noteworthy here. The first is that phonological phenomena responsible for the stem modifications are synchronically inoperative. The second is that these different phenomena are distributed according to a recurrent pattern – which is visible also in the stress position within the regular conjugation: on the one hand 1,2,3sing 3PL on the other 1,2PL (Vincent, 1988).

The most common irregular modifications of preterit stems concern lengthening of final consonant (*venire*), replacing of final consonant with /s/ (*perdere*), and other more complex phenomena that include ablauting of root vowel and deletion/insertion of consonants (*fondere*). Even in the preterit a morphological pattern of stem alternation is detectable, which distinguishes 1,3sing 3PL vs. 2sing 1,2PL (see Table 8).

Even though the majority of alterations were originally phonologically motivated, their present distribution is morphological in nature and paradigmatically governed.

Affixes

Obligatoriness, higher systematicity in mutual relationships, and degree of productivity clearly distinguish inflectional from derivational affixes in fusional languages. Yet the number of productive Italian derivational affixes and the variety of meanings they express are fairly high (Grossmann and Rainer, 2004). Whereas derivation employs both suffixes and prefixes, inflection employs suffixes only. In productive processes the degree of phonological fusion between the stem and derivational affixes is rather low. A higher degree of phonological integration is present instead with interfixes. Some interfixes are semantically void (*congress-<u>u</u>-ale* 'concerning congresses'), the majority express connotative value when combined with evaluative suffixes (*test-<u>ol</u>-ina* 'nice small head'). Words containing interfixes are rather difficult to be segmented. This is because interfixes occupy an intermediate position between the root and the suffix, and generally their morphemic and syllabic boundary does not coincide.

The clear-cut difference between affixes and roots characterizes Italian as a fusional language and distinguishes it from agglutinating languages, which grant affixes greater autonomy (Plungian, 2001). In Italian the coordination between bound elements referring to a same root is marginal (*micro- e macroeconomia* 'micro- and macroeconomy'); moreover, there is no use of a single affix referring to a coordinated group (cf. Spanish *afirmativa o negativa<u>mente</u>* 'affirmatively or negatively'). Another fusional trait of Italian derivation is the signaling through the change of inflectional class of derivatives formed by means of conversion, which belong to the same word class as the base (*banana* FEM, PL *-e* 'banana'-> *banano* MASC, PL *-i* 'banana tree').

Word Order Mobility

The Italian basic word order (SVO) is more flexible than what we encounter in less-fusional languages.

Agreement (in gender and number) provides cohesion among words within phrases, and a certain degree of mobility freedom for the phrases themselves. Instead, single words are less free to move compared to words in more inflecting languages, such as Latin, for instance, (Simone, 1993). Whereas in Latin, the use of cases allows to signal on each word its relational syntactic functions, thus rendering it relatively autonomous within the phrase, in Italian the analytic expression (through prepositions) of syntactic relations demands the proximity and the reciprocal ordering of words within the phrases. Thus, order variation occurs primarily at the level of the reciprocal ordering of phrases (many instances of topicalization, cleft sentences, postverbal subject position). Adjectives stand out from other word classes for a higher degree of movement freedom within the phrases. Even though the unmarked position of the qualifying adjective is postnominal (according to the basic order SVO), this may vary, which at times effects a change in meaning (cf. *famiglie numerose* 'large families' and *numerose famiglie* 'several families').

See also: Arabic as an Introflecting Language; Chinese as an Isolating Language; Finnish as an Agglutinating Language; Italian; Morphological Typology; Morphological Universals; Morphophonemics; Natural Morphology.

Bibliography

D'Achille P & Thornton A M (2003). 'La flessione del nome dall'italiano antico all'italiano moderno.' In Maraschio N & Poggi Salani T (eds.) *Italia linguistica anno mille Italia linguistica anno duemila.* Rome: Bulzoni.

Dressler W U (1985). 'Typological aspects of natural morphology.' *Acta Linguistica Academiae Scientiarum Hungaricae 35,* 51–70.

Dressler W U, Kilani-Schoch M, Spina R & Thornton A M (2003). 'Le classi di coniugazione in italiano e in francese.' In Giacomo-Marcellesi M & Rocchetti A (eds.) *Il verbo italiano.* Rome: Bulzoni.

Greenberg J H (1954). 'A quantitative approach to the morphological typology of language.' In Spencer R F (ed.) *Method and perspective in anthropology.* Minneapolis: University of Minnesota Press.

Grossmann M & Rainer F (eds.) (2004). *La formazione delle parole in italiano.* Tübingen: Niemeyer.

Matthews P H (1970). 'Recent developments in morphology.' In Lyons J (ed.) *New horizons in linguistics.* Harmondsworth: Pelican.

Pirrelli V & Battista M (2000). 'The paradigmatic dimension of stem allomorphy in Italian verb inflection.' *Italian Journal of Linguistics 12,* 307–379.

Plungian V A (2001). 'Agglutination and flection.' In Haspelmath M, König E, Oesterreicher W & Raible W (eds.) *Language typology and language universals.* Berlin, New York: De Gruyter.

Skalička V (1966). 'Ein "typologisches Konstrukt."' In Skalička V (ed.) (1979). *Typologische Studien.* Braunschweig: Vieweg.

Simone R (1993). 'Stabilità e instabilità nei caratteri originali dell'italiano.' In Sobrero A (ed.) *Introduzione all'italiano contemporaneo.* Rome-Bari: Laterza.

Thornton A M (2001). 'Some reflections on gender and inflectional class assignment in Italian.' In Schaner-Wolles C *et al.* (eds.) *Naturally!* Turin: Rosenberg & Sellier.

Vincent N (1988). 'Italian.' In Harris M & Vincent N (eds.) *The Romance languages.* London: Routledge.

Italian Lexicography

A Alonge, University of Perugia, Perugia, Italy

The birth of Italian monolingual lexicography can be traced back to the end of the 15th century, when the first lists of vernacular words were compiled. The first printed monolingual inventory, *Le tre fontane* by Niccolò Liburnio, was published in Venice some time later, in 1526. Partly following Bembo's proposals with respect to the *querelle de la langue* (according to which Petrarch's language should have been the model for poetry and Boccaccio's language for prose), Liburnio lists and defines terms taken from the works of Dante, Petrarch, and Boccaccio. Other dictionaries published in those years had the same purpose, to encode and define the lexicon of the 14th-century Tuscan classics, to be imitated by those aspiring to linguistic elegance: e.g., Lucilio Minerbi's *Vocabulario* (Venice, 1535); Francesco Alunno's *Le Osservazioni sopra il Petrarca* (Venice, 1539; 1550), and *Le ricchezze della lingua volgare sopra il Boccaccio* (1543).

The *Vocabulario di cinquemila Vocabuli Toschi* by Fabricio Luna (Naples, 1536) has been described as the first true Italian dictionary, ordered alphabetically and listing 5000 terms taken from the works of various authors of different cultural and geographical origins, but also from contemporary usage. Alberto Acarisio's *Vocabolario, grammatica, et ortographia de la lingua volgare* (Cento, 1543) is interesting both for the richness of data encoded for each lemma (compared to other works of the time, Acarisio's work pays more attention to scientific terminology

Agreement (in gender and number) provides cohesion among words within phrases, and a certain degree of mobility freedom for the phrases themselves. Instead, single words are less free to move compared to words in more inflecting languages, such as Latin, for instance, (Simone, 1993). Whereas in Latin, the use of cases allows to signal on each word its relational syntactic functions, thus rendering it relatively autonomous within the phrase, in Italian the analytic expression (through prepositions) of syntactic relations demands the proximity and the reciprocal ordering of words within the phrases. Thus, order variation occurs primarily at the level of the reciprocal ordering of phrases (many instances of topicalization, cleft sentences, postverbal subject position). Adjectives stand out from other word classes for a higher degree of movement freedom within the phrases. Even though the unmarked position of the qualifying adjective is postnominal (according to the basic order SVO), this may vary, which at times effects a change in meaning (cf. *famiglie numerose* 'large families' and *numerose famiglie* 'several families').

See also: Arabic as an Introflecting Language; Chinese as an Isolating Language; Finnish as an Agglutinating Language; Italian; Morphological Typology; Morphological Universals; Morphophonemics; Natural Morphology.

Bibliography

D'Achille P & Thornton A M (2003). 'La flessione del nome dall'italiano antico all'italiano moderno.' In Maraschio N & Poggi Salani T (eds.) *Italia linguistica anno mille Italia linguistica anno duemila.* Rome: Bulzoni.

Dressler W U (1985). 'Typological aspects of natural morphology.' *Acta Linguistica Academiae Scientiarum Hungaricae 35,* 51–70.

Dressler W U, Kilani-Schoch M, Spina R & Thornton A M (2003). 'Le classi di coniugazione in italiano e in francese.' In Giacomo-Marcellesi M & Rocchetti A (eds.) *Il verbo italiano.* Rome: Bulzoni.

Greenberg J H (1954). 'A quantitative approach to the morphological typology of language.' In Spencer R F (ed.) *Method and perspective in anthropology.* Minneapolis: University of Minnesota Press.

Grossmann M & Rainer F (eds.) (2004). *La formazione delle parole in italiano.* Tübingen: Niemeyer.

Matthews P H (1970). 'Recent developments in morphology.' In Lyons J (ed.) *New horizons in linguistics.* Harmondsworth: Pelican.

Pirrelli V & Battista M (2000). 'The paradigmatic dimension of stem allomorphy in Italian verb inflection.' *Italian Journal of Linguistics 12,* 307–379.

Plungian V A (2001). 'Agglutination and flection.' In Haspelmath M, König E, Oesterreicher W & Raible W (eds.) *Language typology and language universals.* Berlin, New York: De Gruyter.

Skalička V (1966). 'Ein "typologisches Konstrukt."' In Skalička V (ed.) (1979). *Typologische Studien.* Braunschweig: Vieweg.

Simone R (1993). 'Stabilità e instabilità nei caratteri originali dell'italiano.' In Sobrero A (ed.) *Introduzione all'italiano contemporaneo.* Rome-Bari: Laterza.

Thornton A M (2001). 'Some reflections on gender and inflectional class assignment in Italian.' In Schaner-Wolles C *et al.* (eds.) *Naturally!* Turin: Rosenberg & Sellier.

Vincent N (1988). 'Italian.' In Harris M & Vincent N (eds.) *The Romance languages.* London: Routledge.

Italian Lexicography

A Alonge, University of Perugia, Perugia, Italy

The birth of Italian monolingual lexicography can be traced back to the end of the 15th century, when the first lists of vernacular words were compiled. The first printed monolingual inventory, *Le tre fontane* by Niccolò Liburnio, was published in Venice some time later, in 1526. Partly following Bembo's proposals with respect to the *querelle de la langue* (according to which Petrarch's language should have been the model for poetry and Boccaccio's language for prose), Liburnio lists and defines terms taken from the works of Dante, Petrarch, and Boccaccio. Other dictionaries published in those years had the same purpose, to encode and define the lexicon of the 14th-century Tuscan classics, to be imitated by those aspiring to linguistic elegance: e.g., Lucilio Minerbi's *Vocabulario* (Venice, 1535); Francesco Alunno's *Le Osservazioni sopra il Petrarca* (Venice, 1539; 1550), and *Le ricchezze della lingua volgare sopra il Boccaccio* (1543).

The *Vocabulario di cinquemila Vocabuli Toschi* by Fabricio Luna (Naples, 1536) has been described as the first true Italian dictionary, ordered alphabetically and listing 5000 terms taken from the works of various authors of different cultural and geographical origins, but also from contemporary usage. Alberto Acarisio's *Vocabolario, grammatica, et ortographia de la lingua volgare* (Cento, 1543) is interesting both for the richness of data encoded for each lemma (compared to other works of the time, Acarisio's work pays more attention to scientific terminology

and to phraseology) and for an attempt to encode many lemmas that are not found in the works of the 14th-century classics. Other prominent lexicographical works of the time are Alunno's *Fabrica del mondo* (Venice, 1548), which is divided into 10 sections and may be considered the first methodical Italian dictionary, and Giacomo Pergamini's *Memoriale della lingua*, first published in Venice in 1602 and then reprinted various times.

The first *Vocabolario degli accademici della Crusca* was published in Venice in 1612 and became a model for monolingual European lexicography. It conformed to new and scientific criteria: orthographical and grammatical norms were stated, more coherent modes of defining entries were used, and various standardizing expedients were adopted. Again, the main aim of the work was to indicate the best linguistic norm, by referring to the three main 14th-century writers, but also to other texts written in Florence in the same period, following proposals put forward by Leonardo Salviati. The *Vocabolario* was enlarged and republished many times until 1923 and, although criticized because of the limits it imposed on contemporary language use, it contributed substantially to the definition and diffusion of a national language. A new project, initiated in 1965, aims at historical documentation and is now realized by the *Opera del Vocabolario Italiano*, an institute of the C.N.R. (*Consiglio Nazionale delle Ricerche*).

Various dictionaries were compiled in the 17th and 18th centuries aimed at contesting the primacy of the Tuscan language, while other works followed fundamental principles which had inspired the *Crusca* dictionary, indicating possible additions and innovations to it.

In the 19th century, while the *querelle de la langue* was still engaging personages like Alessandro Manzoni and Graziadio Isaia Ascoli, new dictionaries were produced that included new technical terms, referred to new sources and took into account current colloquial usage. However, the most important work of the time is the *Dizionario della lingua italiana* (Torino, 1861–1879) by Niccolò Tommaseo and Bernardo Bellini: notwithstanding some limitations of the work (e.g., etymologies are often ill-founded), the richness of examples still makes it a useful lexicographical tool.

Among the many dictionaries of Italian published from the beginning of the 20th century up to now, we find both large scholarly dictionaries and more practical ones. The *Grande Dizionario della lingua italiana* (1961–2002), founded by Salvatore Battaglia, is an unsurpassed historical dictionary (UTET, Turin) comprising 21 volumes. The *Lessico Universale Italiano* (Istituto dell'Enciclopedia Italiana, Rome, 1968–1981), in 24 volumes, has qualities of both an encyclopedia and a dictionary. Another great dictionary is the *Grande dizionario d'uso della lingua italiana* directed by Tullio De Mauro and containing 263 000 lemmas: first published in 1999 (UTET), it was updated in 2003, when the seventh volume was published. Among the many practical dictionaries, just two examples may be mentioned: Nicola Zingarelli's *Vocabolario della lingua Italiana*, first published in 1917, then reprinted and updated many times up to the present day, most recently by Miro Dogliotti and Luigi Rosiello (Zanichelli, Milan, 2004); and the *Dizionario della lingua italiana* by Giacomo Devoto and Gian Carlo Oli, first printed in 1967 and also republished many times: the 2004–2005 edition is by Luca Serianni and Maurizio Trifone (Le Monnier, Florence). While electronic versions of these and other contemporary dictionaries are also published, pioneering work has been carried out (notably at the ILC-CNR [*Istituto di Linguistica Computazionale-Consiglio Nazionale delle Ricerche*], Pisa) to produce online lexical resources, first of all the *Dizionario-macchina dell'italiano*. Another lexical online resource recently developed (*ILC*, 1996–2001) is *ItalWordNet*, a thesaurus-like resource, basically structured in the same way as *WordNet,* a system whose design is inspired by current psycholinguistic theories of human lexical memory, developed at the Cognitive Science Laboratory at Princeton University in the United States.

Beside monolingual dictionaries, other lexicographical works are worthy of mention: etymological dictionaries (e.g., the one by M. Cortelazzo and P. Zolli, recently republished in 1999 by Zanichelli); frequency lexicons (LIF [*Lessico italiano di frequenza]*, by U. Bortolini, C. Tagliavini, and A. Zampolli, for Garzanti-IBM, Milan, 1972; LIP [*Lessico di frequenza dell'italiano parlato],* T. De Mauro, F. Mancini, M. Vedovelli, M. Voghera, for Etas, Milan, 1993); regional dictionaries (a notable example is the *Vocabolario Siciliano*, founded by G. Piccitto and subsequently directed by G. Tropea, 1977–2002).

See also: Academies: Dictionaries and Standards; Ascoli, Graziadio Isaia (1829–1907); Computational Lexicons and Dictionaries; Computers in Lexicography; Dictionaries and Encyclopedias: Relationship; Italian; Italy: Language Situation; Thesauruses; WordNet(s).

Bibliography

Della Valle V (1994). 'La lessicografia.' In Serianni L & Trifone P (eds.) *Storia della lingua italiana.* Torino: Einaudi. 25–91.

Massariello Merzagora G (1983). *La lessicografia.* Bologna: Zanichelli.

Migliorini B (1961). *Che cos'è un vocabolario?* Firenze: Le Monnier.
Poggi Salani T (1982). 'Venticinque anni di lessicografia italiana delle origini (leggere, scrivere e "politamente parlare"): note sull'idea di lingua.' *Historiographia Linguistica IX*. 265–297.
Rossebastiano Bart A (1986). 'Alle origini della lessicografia italiana.' In Buridant C (ed.) *Lexique 4: La lexicographie au Moyen Age*. Lille: UMR SILEX. 113–156.
Tancke G (1984). *Die italienischen Wörterbücher von den Anfängen bis zum Erscheinen des 'Vocabolario degli Accademici della Crusca' (1612). Bestandsaufnahme und Analyse*. Tübingen: Niemeyer.
Zampolli A, Calzolari N & Cignoni L (eds.) (2003). *Linguistica Computazionale 16–17. Special Issue on Computational Linguistics in Pisa*. Pisa: Giardini. 745–792.

Italy: Language Situation

A L Lepschy
G Lepschy

This article is reproduced from the previous editions, volume 4, pp. 1792–1793, © 1994, Elsevier Ltd.

It is traditionally taken for granted that Italy is basically a monolingual state, and that Italian is the national/official language, used not only in literature but also in government, administration, newspapers, and of course as the medium of instruction in schools and universities.

The reality is different. It is useful to consider two distinct sets of facts. One concerns the existence of so-called minority languages within the Italian territory. There are enclaves in which foreign languages are spoken; this may result from an earlier immigration of communities that preserved their language through the centuries – as is the case in villages in central and southern Italy, where the inhabitants speak Albanian (ca. 100 000), Greek (ca. 30 000), Catalan in Alghero (ca. 15 000), and Serbo–Croatian (ca. 3000); or it may result from the inclusion within the Italian boundaries of populations speaking foreign languages – as in the case of German in South Tyrol (ca. 280 000), Occitan and Franco-Provençal in Piedmont (ca. 115 000), and Slovene in the Veneto (ca. 53 000).

The second set of facts concerns the use of so-called Italian dialects. These are separate and, on the whole mutually unintelligible, Romance languages, classifiable into 15 or so regional groups (Ligurian, Piedmontese, Lombard, Venetian, Tuscan, etc.); they are cognates (rather than varieties) of literary Italian, which is itself based on Tuscan. Many linguists consider Sardinian, Ladin, and particularly Friulian, as separate languages. Here the situation is complicated by the difficulty of establishing how widespread the use of dialect is. It was calculated that at the time of unification only a small minority (suggested figures range between 2.5–10%) of the inhabitants of the new state knew Italian. It appears that in the 1980s, about half of all Italians still used a dialect at home and with friends (this applies particularly to southern Italy and the Veneto). The novelty is that even dialect speakers are mostly bilingual with Italian. Although traditionally the dialect was used in all circumstances (some speakers employing it even in the context of high culture), diglossia seems to prevail, Italian being the high variety (used in public life, in official contexts, and for elevated topics), and dialect the low variety (used in the home and for personal, everyday exchanges).

The status of minority languages was, until World War II, mostly unrecognized. Successively the regions of Trentino–Alto Adige and of the Valle d'Aosta have obtained full recognition for the use of German and French, respectively, in education and in public life. This also applies, to a lesser extent, to Sardinia and Friuli–Venezia Giulia. Political pressure is mounting to extend similar recognition to other 'foreign' minority languages, and even to some of the regional dialects, such as Lombard or Venetian.

Complex linguistic problems are posed by new waves of immigration, mainly from Africa, in the late 1980s and early 1990s. As for foreign languages studied in schools, French, which was dominant until World War II, has since lost its primacy to English.

Bibliography

De Mauro T (1976). *Ctoria linguistica dell'Italia unita*. Bari: Laterza.
Lepschy A L & Lepschy G (1991). *The Italian language today*. London: Routledge.

Migliorini B (1961). *Che cos'è un vocabolario?* Firenze: Le Monnier.
Poggi Salani T (1982). 'Venticinque anni di lessicografia italiana delle origini (leggere, scrivere e "politamente parlare"): note sull'idea di lingua.' *Historiographia Linguistica IX*. 265–297.
Rossebastiano Bart A (1986). 'Alle origini della lessicografia italiana.' In Buridant C (ed.) *Lexique 4: La lexicographie au Moyen Age*. Lille: UMR SILEX. 113–156.
Tancke G (1984). *Die italienischen Wörterbücher von den Anfängen bis zum Erscheinen des 'Vocabolario degli Accademici della Crusca' (1612). Bestandsaufnahme und Analyse*. Tübingen: Niemeyer.
Zampolli A, Calzolari N & Cignoni L (eds.) (2003). *Linguistica Computazionale 16–17. Special Issue on Computational Linguistics in Pisa*. Pisa: Giardini. 745–792.

Italy: Language Situation

A L Lepschy
G Lepschy

This article is reproduced from the previous editions, volume 4, pp. 1792–1793, © 1994, Elsevier Ltd.

It is traditionally taken for granted that Italy is basically a monolingual state, and that Italian is the national/official language, used not only in literature but also in government, administration, newspapers, and of course as the medium of instruction in schools and universities.

The reality is different. It is useful to consider two distinct sets of facts. One concerns the existence of so-called minority languages within the Italian territory. There are enclaves in which foreign languages are spoken; this may result from an earlier immigration of communities that preserved their language through the centuries – as is the case in villages in central and southern Italy, where the inhabitants speak Albanian (ca. 100 000), Greek (ca. 30 000), Catalan in Alghero (ca. 15 000), and Serbo–Croatian (ca. 3000); or it may result from the inclusion within the Italian boundaries of populations speaking foreign languages – as in the case of German in South Tyrol (ca. 280 000), Occitan and Franco-Provençal in Piedmont (ca. 115 000), and Slovene in the Veneto (ca. 53 000).

The second set of facts concerns the use of so-called Italian dialects. These are separate and, on the whole mutually unintelligible, Romance languages, classifiable into 15 or so regional groups (Ligurian, Piedmontese, Lombard, Venetian, Tuscan, etc.); they are cognates (rather than varieties) of literary Italian, which is itself based on Tuscan. Many linguists consider Sardinian, Ladin, and particularly Friulian, as separate languages. Here the situation is complicated by the difficulty of establishing how widespread the use of dialect is. It was calculated that at the time of unification only a small minority (suggested figures range between 2.5–10%) of the inhabitants of the new state knew Italian. It appears that in the 1980s, about half of all Italians still used a dialect at home and with friends (this applies particularly to southern Italy and the Veneto). The novelty is that even dialect speakers are mostly bilingual with Italian. Although traditionally the dialect was used in all circumstances (some speakers employing it even in the context of high culture), diglossia seems to prevail, Italian being the high variety (used in public life, in official contexts, and for elevated topics), and dialect the low variety (used in the home and for personal, everyday exchanges).

The status of minority languages was, until World War II, mostly unrecognized. Successively the regions of Trentino–Alto Adige and of the Valle d'Aosta have obtained full recognition for the use of German and French, respectively, in education and in public life. This also applies, to a lesser extent, to Sardinia and Friuli–Venezia Giulia. Political pressure is mounting to extend similar recognition to other 'foreign' minority languages, and even to some of the regional dialects, such as Lombard or Venetian.

Complex linguistic problems are posed by new waves of immigration, mainly from Africa, in the late 1980s and early 1990s. As for foreign languages studied in schools, French, which was dominant until World War II, has since lost its primacy to English.

Bibliography

De Mauro T (1976). *Ctoria linguistica dell'Italia unita*. Bari: Laterza.
Lepschy A L & Lepschy G (1991). *The Italian language today*. London: Routledge.

Ivory Coast: Language Situation

F Ahoua, Université de Cocody, Abidjan, Côte d'Ivoire

Ivory Coast, officially named Côte d'Ivoire, is situated in West Africa, approximately between Latitude 8° 00′ N and Longitude 5° 00′W. The country has an area of 322 460 km^2 and a population density of 50 inhabitants/km^2. The country is bordered on the south by the Atlantic Ocean, on the west by Guinea and Liberia, on the north by Mali and Burkina Faso.

The population is currently estimated at 17 327 724. The official census of 1998 (*Récensement Général de la Population et de l'Habitat*) estimated the literacy rate at 45%. The percentage of males is about 57% while that of females is about 43% (2003 estimate). The percentage of foreign immigrants is believed to be close to 26%, which has consequences for the migrant languages spoken in Ivory Coast (mainly Moore [Mòoré], Fulfulde, Soninke, Maninkan [Maninka], Bambara [Bamanankan], Krahn, Glio [Glio-Oubi]).

Number of Languages and Language Families in Côte d'Ivoire

The French colonial borders – that were established from 1893 to 1960 and definitely remained – did not take into consideration the actual ethnic and language boundaries and used political and geographic criteria. This has resulted in the splitting of the language groups that lie along the borders (see also Kropp-Dakubu 2000a,b,c).

For instance, the languages Anyi and Nzema are spoken both in Ghana and Ivory Coast, across the political borders of today. Speakers of the Senoufo and Manding language families stretch beyond the borders of Ivory Coast, into Mali, Sierra Leone, Guinea, and Burkina Faso. Dan and Guere are spoken in Ivory Coast and Liberia as well. The languages were spoken in these areas long before the colonial boundaries were established. This situation is different from most recent cases of immigration, involving the use of a language in the already established borders by the French and British.

The heterogeneity and the relatively large but so far often inaccurate number of languages of Ivory Coast raise the question of the criteria of language identification (as opposed to dialects), mutual intelligibility, genetic classification, relationship between social status and levels of usage, and political or cultural importance of the languages. The issue impels one to clarify which criteria are necessary or sufficient, and in many cases go beyond mere linguistic criteria. For example, earlier works on Ivorian languages (Delafosse, 1904) equated language with ethnicity, while even more recent sources often (Ethnologue, 1988) do not distinguish between local languages and recent foreign languages.

The following list presents the indigenous languages of Ivory Coast in alphabetical order based on most recent linguistic criteria:

- Abbey (Abé)
- Abidji
- Abron
- Abure
- Adiukru (Adioukrou)
- Aizi, Aproumu
- Aizi, Mobumrin
- Aizi, Tiagbamrin
- Alladian
- Anyin
- Attié
- Avikam
- Baule (Baoulé)
- Beng
- Bété, Daloa
- Bété, Guiberoua
- Beti (Eotile)
- Dan (Yacuba)
- Dida, Lakota (Dieko, Gabo, Satro, Guébie, Brabori, Ziki)
- Dida, Yocoboué
- Dioula (Jula)
- Ebrié (Cama)
- Ega (Diés, Egwa)
- Esuma (Essouma)
- Gagu (Gagou, Gban)
- Godié
- Guéré, Central (Wé, Southern)
- Guro (Kweni)
- Kodia (Kwadia, Kwadya)
- Kouya (Kowya, Sokya, Kuya)
- Krobu (Krobou)
- Krumen
- Kulango, Bondoukou
- Kulango, Bouna
- Mwan (Muan, Mona, Mouan, Muana, Mwa)
- Neyo (Gwibwen, Towi)
- Nyabwa (Nyaboa, Niaba,Nyédébwa)
- Nzema (Appolo, Nzima)
- Senoufo, Cebaara (Senari)
- Senoufo, Djimini
- Senoufo, Niarafolo-Niafolo (Senoufo, Nyarafolo)
- Senoufo, Palaka
- Senoufo, Shempire
- Senoufo, Tagwana

- Tééń
- Toura
- Wan
- Wané
- Wobe (Wɛ̃, Northern)
- Yaouré

Despite the large diversity of the ethnic groups, it is possible to group the Ivorian languages into four greater families by using linguistic criteria: Mande, Kru, Akan, Voltaic or Gur (cf. map below, Figure 1). The name 'Voltaic' has long been used in the francophone area but is being abandoned for the name 'Gur,' which means 'language' (Gurma, Gurunsi).

1. The Mande group is located in the northeast of Côte d'Ivoire, and covers the Manding (Delafosse's Malinke, Bambara, Dioula) and the southern Mande. In the western central part of the country live the Dan in a mountaneous area, surrounding the city of Man (cf. Braconnier *et al*., 1983; Tera, 1986).
2. The Kru group: The central-western and south-western areas are occupied by the Kru or Maguwe, of which the major ethnic group is the Bete.
3. The Gur group contains Senufo as a set of diverse languages including Lobi, Tagbana, Cebaara, and is geographically located in the north.
4. The Kwa group covers the east, center, and south-east of Ivory Coast. The most numerous sub-group is the Baule language with about 60 dialects. Sixteen languages compose the Kwa group in the Ivory Coast: Baule, Agni (Anyin), Abron, the lagoon languages (Ebrié, Abouré, Esuma, Adiukrou, Abidji, Avikam, Alladian, Eotile, Abbey, Mbatto or Ngula (Mbato or Gwo), Nzema, Abidi (Abidji)), and three isolates: Ega, Attie, Krobu.

An official census conducted in 1998 provides the following percentages of the individual ethnic groups: Akan 42.1%, Gur 17.6%, Northern Mande 16.5%, Kru 11%, Southern Mande 10%, others 2.8% (including only 130 000 Lebanese and 14 000 French). If that census is to take into account all the African and other ethnic groups, one has to revise the percentages by computing the approximate 26% that constitute one-third of the total population of Ivory Coast.

The non-Ivorian population includes people from Malba Birifor, Bissa (63 000), Dogoso, Maasina Fulfulde (1200), Glaro-Twabo, Eastern Karaboro (5000), Khe, Mòoré, Nafaanra, Sìcìté Sénoufo, Siamou, Southern Toussian, Turka, Vietnam, Burkina Faso (3 000 000), Mali (754 000), Guinea (238 000), Liberia (200 000) and others (345 000) (1993). Ghanaians approach almost 500 000, Europeans almost 20 000 (the number includes French, binationals, Germans, and Belgians). Syro-Lebanese are estimated at 100 000.

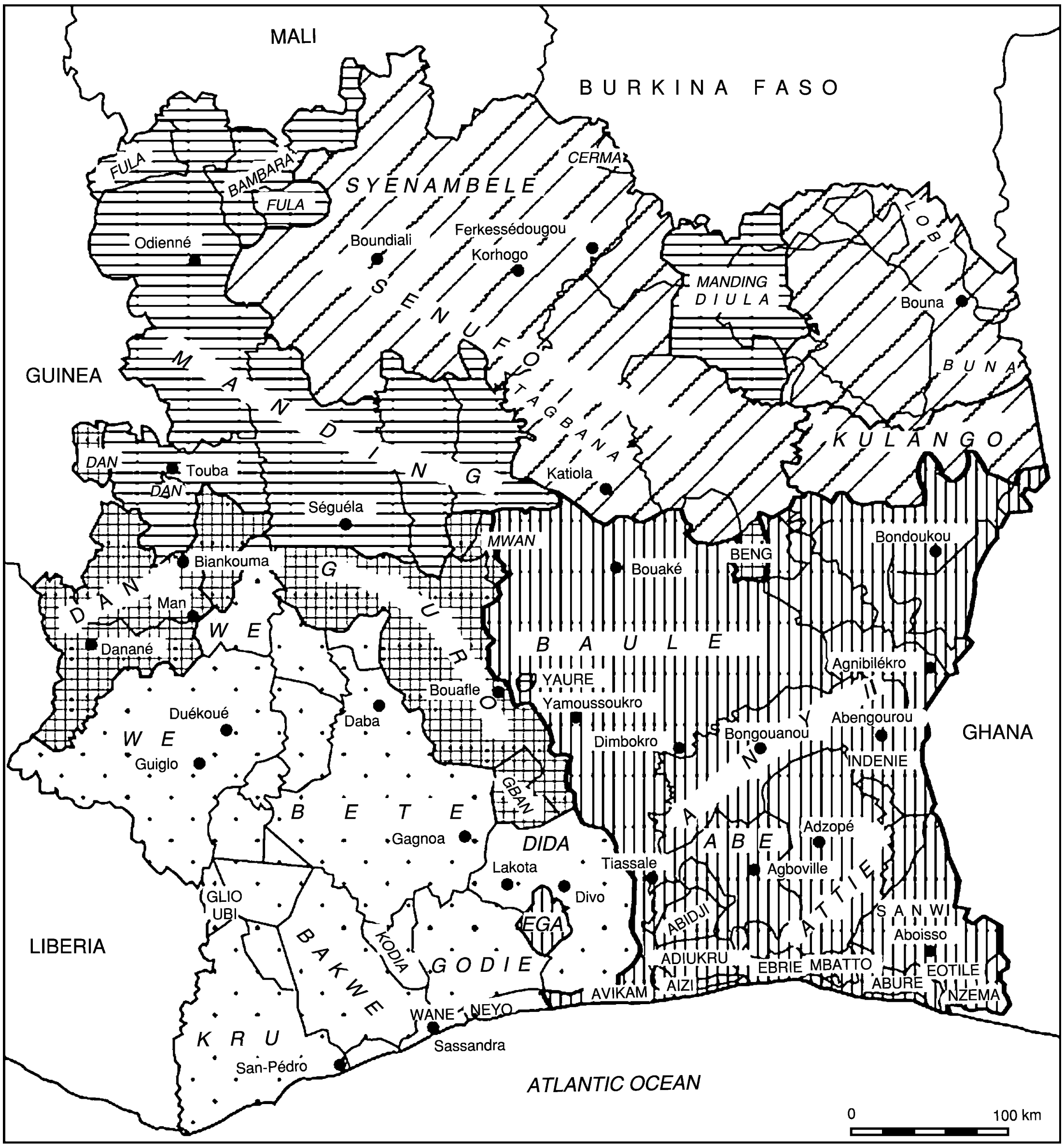
The geographic distribution of the ethnic groups is as follows in the north Senoufos (9.7%), Dioulas (3.4%), Malinkes (8.5%); in the center Guros (3.6%) and Baules (6.6%); in the west the Dans or Yacubas (5.9%), Betes (5.7%), Gueres (3.4%), Wobes (1.7%), Krumen-Didas-Bakwes-Godies (2.1%); in the south, the Ebriés (Cama) -Mbatos (Gwas) (0.7%), Adiukrus-Abidjis-Alladians-Avikams (1%), and Nzemas, Abbeys (1.4%).

The largest individual (first) language in Ivory Coast is Baule, because the Senufos use different languages. In practice, the language that is most used as a second language in markets and for commercial, cultural, and religious activities is Dioula. The list of the languages in Table 1 is not comprehensive, but includes the 17 most widely distributed languages. Though the list indicates 1.2% native Dioulas, the number of users may be about 14.5%. Investigations made on conversations in the markets estimate that 88% of the verbal exchanges in such environments are performed in Dioula.

Multilingualism and Diglossia

The national or official language, the *lingua franca* is French. French still remains the official language, widely used in schools, sometimes at different school levels, and is the obligatory language for administration and the major language of the media. Because of the contact situation, a local indigenous French has emerged, including new variants, one of which is called 'nouchi,' a slang or a secret language derived from French and the Ivorian languages, used as in daily communication and in popular songs.

Multilingualism is the most common factor among all the minority languages. The issue is not how perfect the second language is under control but how often it is used. Second languages are frequently used on a daily basis. For instance, all the Manding (formerly Malinke) use Dioula as a second language in religious, commercial, and cultural activities. This situation is valid for all the Senufos. The Ega use Dida in everyday life, but their own Ega language in the family, though the situation is more complex, because exogamy favors marriage with the Dida women, so that Ega children grow up with Dida. All the Eotile speak Anyi, Abure, and Nzemas at the respective ethnic borders where the Eotile villages are located in enclaves. All the Aizi perfectly speak Adiukru, a major, more commercially and culturally powerful



Source : Atlas of C. I., plate B 2

Conception: AHOUAF-realisation: SANOGOS.IGT-LATIG janvier 2005

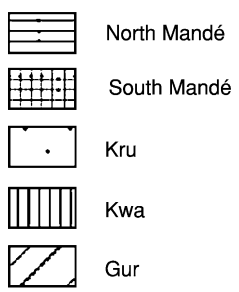


Figure 1 Language use in Ivory Coast.

Table 1 Language use in Ivory Coast

Language	*Number of Speakers*	*Percentage*	*Linguistic Affiliation*
1. Baule	2 130 000	14.8%	Kwa
2. Senufo	1 245 500	8.7%	Gur
3. Dan	800 000	5.5%	Mande
4. Anyi	610 000	4.2%	Kwa
5. Attie	381 000	2.6%	Kwa
6. Guere	317 688	2.2%	Kru
7. Bete	280 000	1.9%	Kru
8. Dioula	179 000	1.2%	Mande
9. Abbey	170 000	1.1%	Kwa
10. Mahou	169 100	1.1%	Mande
11. Wobe	156 300	1.0%	Kru
12. Lobi	155 800	1.0%	Gur
13. Kulango	142 000	0.9%	Kru
14. Brong (Abron)	131 700	0.9%	Kwa
15. Dida	101 600	0.7%	Kru
16. Adiukru	100 000	0.6%	Kwa
17. Soninke	100 000	0.6%	Mande

ethnic group. The Krobu in general speak Adiukru and Abbey very fluently, and sometimes Baule. The Mbatto generally speak Attie or Abure as second languages.

The Wan speak Guro and Baule-Kode and the Nguin speak Baule as a second language. The Mau generally use Dioula as a second language but only older Mau people speak Dan fluently, in specific cases of trade with the Dans. The Yaure speak Guro and Baule as second languages, respectively at the Guro and Baule borders; the Gagu use Baule and Bete as second languages, respectively at the Baule and Guro borders. The Brong and Kulango use each other languages as second languages, and the Tagbanas are generally bilingual with Baule.

In all cases of multilingualism in Ivory Coast, specific situations determine whether the speakers prefer one type of language or the other, suggesting a strong case for diglossia. The typical language that is used for religious and commercial cases is Dioula, at the market (about 70%), and secondly Ivorian French and Baule (30%) in proportion with dominance of Dioula in trade activities. French is definitely used for administration and the local native languages are used at home.

Ivorian French and Argots

It is estimated that 36% of the population up to six years old use a variant of French called 'Français Populaire Ivoirien' (cf. Djite, 1989a,b). More than 1 million of the illiterate population use a pidgin French called 'Français de Moussa' or 'Francais de Dago,' which is more or less similar to 'Français de petit nègre,' difficult to understand for non-Ivorian Francophones.

A number of researchers and most recent studies have established a Standard Ivorian French used in the official newspapers and in the public media (radio and television) and by students (cf. Boutin, 2002; Hattiger, 1981; Lafage, 1989).

Though Kouadio (2000) suggests that there are three types of variants, a high, middle, and lower class level, in practice, there may be also three registers: formal, informal, and popular or pidgin that are used by all classes, according to the situation and the speakers. The formal register is used in the public media, schools and universities, the second in informal situations, and the last by illiterate people.

Variants of French used by illiterates, called Moussa or Dago French (pidgin types according to Kokora 1980, 1983a,b) as mentioned above, are quite well known and have been often stigmatized in the newspapers or on the TV and radio.

Nouchi is a secret language that has been made popular, thanks to well known music groups (Zuglu) and local reggae musicians (Serges Kassy).

Language Policy and Use

There has been no clear policy for the development of local languages. Many projects have set French as the priority standard language. The reason is that, contrary to other West African Francophone countries, there is no national language such as Wolof in Senegal, Bambara in Mali, Moore in Burkina Faso, Gen-Mina in Togo, Fon (Fon-Gbe) in Benin, and Akan in Ghana. Though the Baule and Anyi block are the dominant languages, they are not the dominant trade languages as is Dioula, which is a favorite for a national language because of its widespread use all over the country (cf. Djite, 1988; Tera 1986).

Nevertheless, there were attempts to introduce local languages, such as Dan, Cebaara, and Attie, until the CE1 or CM2 (that is, from Grade I to Grade 5) in some very restricted projects. But although such projects proved to be successful for the teaching of sciences and Maths, they never acquired a national status due to their often detrimental effects on the acquisition of spoken French.

Most descriptive work on Ivorian languages, especially with respect to the Atlases (cf. Hérault 1983 on Kwa, Marchese 1983 on Kru, Mensah and Tchagbale 1983 on Gur, Halaoui *et al.* 1983 on South-Mande, Braconnier *et al.* 1983 on North-Mande) were carried out by the Institute of Applied Linguistics which was created by Law no. 66–375 of September 8, 1966. Further additions regarding teaching reforms were introduced in 1977 and extended by Law no. 96–696 of

September 7, 1995. The closest collaborator for the Atlases was the Summer Institute of Linguistics.

The Bible Society has also contributed to translation of the Bibles into the following languages: Abbey, Abidji, Adiukru, Alladian, Anyi, Attie, Avikam, Baule, Bete, Birifior, southern Bissa, Cebaara, Cerma, Dan, Dida, Ebrie, Fula, Macina (Fulfulde, Massina), Gban, Guro, Guere, Diula, Karaboro, Kulango, Krumen, Southern Lobiri, Nyabuwa, Nzema, Senufo, Djimini, Tagbana, Teen, Tura and Wobe.

As reported by Kouadio (2000), 20 hours out of 168 hours per week of radio broadcasting are given in 14 languages, that is: Abbey, Adiukru, Attie, Baule, Bete, Dida, Dioula, Ebrie, Guere, Krumen, Kulange (Kulango, Bondoukou), Senufo, Dan, and Moore. The National Television broadcasts 3 hours out of 81 per week in local languages.

A number of Ivorian languages are isolated among different other larger language groups, for example, Beng (Ngain, Nguin, Gan) with 17 000 speakers, Aizi Mobumri, Wan (Mona), and Gagu, Yaure. Some of these languages are endangered, and documentation on them would benefit the linguistic and the speaker community. Two examples of endangered languages in Ivory Coast are Mbre (Bre, Bere, Pré Pisia), a recently discovered language of unknown affiliation in central Ivory Coast spoken by about 200 speakers the Bondosso and Nitanbo villages, for which only sketchy descriptions exist, and the last remaining dialect of Eotile (Beti, Betibe), a Kwa language whose other dialects are extinct, but whose remaining dialect is used by a fishing community on Aby Lagoon.

See also: Akan; Burkina Faso: Language Situation; Ghana: Language Situation; Gur Languages; Kru Languages; Kwa Languages; Mali: Language Situation; Mande Languages; Niger-Congo Languages.

Language Maps (Appendix): Map 6.

Bibliography

Boutin A B (2002). Description de la Variation: Etudes transformationnelles des Phrases du Francais de Côte d'Ivoire. Ph.D. Diss: Grenoble 3.

Braconnier C, Maire J & Tera K (1983). *Etudes sur le Mandingue de Côte d'Ivoire.* ACCT: Institut de Linguistique Appliquée.

Delafosse M (1904). *Vocabulaires comparatifs de plus de 60 langues ou dialectes parlés à la Côte d'Ivoire.* Paris: Ernest Leroux.

Djite P G (1988). 'The spread of Dyula and popular French in Côte d'Ivoire: implications for language policy.' *Language Problem and Language Planning 12(3),* 213–225.

Djité P G (1989a). 'French in the Ivory Coast.' *The French Review 62(3),* 494–504.

Djité P G (1989b). 'French in the Ivory Coast: a process of nativization.' *Washington Academy of Sciences 79(1),* 1–8.

Grimes J & Grimes B F (eds.) (1988). *Ethnologue. Vol 1: Languages of the World.* Dallas: SIL.

Halaoui N, Kalilou T & Trabi M (1983). *Atlas des langues Mandé-Sud de Côte d'Ivoire.* ACCT: Institut de Linguistique Appliquée.

Hattiger J-L (1981). *Morpho-syntaxe du groupe nominal dans un corpus de français populaire ivoirien.* Université de Strasbourg: Thèse de doctorat 3ème cycle.

Hérault G (1983). Atlas *des langues Kwa de Côte d'Ivoire 1 and 2.* ACCT: Institut de Linguistique Appliquée.

Kokora P D (1980a). 'Trois principes présidant à la formation du lexique en "pidgin français" de Côte-d'lvoire.' *BOFCAN 1,* 39–40.

Kokora P D (1983a). 'Language contact: the case of Ivory Coast.' *CIRL (Abidjan) 13,* 148–172.

Kokora P D (1983b). 'Situation sociolinguistique en Côte-d'lvoire et emprise du français: les variétés de celui-ci et les attitudes langagières qu'elles inscitent chez les locuteurs.' *Cahiers Ivoiriens de Recherche Linguistique 13,* 148–172.

Kouadio N'guessan J (2000). 'Ivory Coast: Language Situation.' In Asher R E & Simpson J M Y (eds.) *The encyclopedia of language and linguistics,* vol. 4. Oxford: Pergamon Press. 1793–1796.

Kropp-Dakubu M E (2000a). 'Akan.' In Asher R E & Simpson J M Y (eds.) *The encyclopedia of language and linguistics,* vol. 1. Oxford: Pergamon Press. 63–64.

Kropp-Dakubu M E (2000b). 'Ghana: Language Situation.' In Asher R E & Simpson J M Y (eds.) *The encyclopedia of language and linguistics,* vol 3. Oxford: Pergamon Press. 1434–1435.

Kropp-Dakubu M E (2000c). 'Kwa Languages.' In Asher R E & Simpson J M Y (eds.) *The encyclopedia of language and linguistics,* vol 4. Oxford: Pergamon Press. 1882–1883.

Lafage S (1989). 'Rôle et place du français populaire dans le continuum langues africaines/français de Côte d'Ivoire.' *Le Français Moderne 47,* 208–219.

Marchese L (1983). *Atlas Linguistique Kru.* ACCT: Institut de Linguistique Appliquée.

Mensah E N A & Tchagbale Z (1983). *Atlas des langues Gur de Côte d'Ivoire.* ACCT: Institut de Linguistique Appliquée.

Tera K (1986). 'Le Dioula véhiculaire de Côte d'Ivoire: expansion et developpement.' *Cahiers Ivoiriens de Recherche Linguistique 20,* 5–32.

Iyer, L. Vishwanatha Ramaswami (1895–1948)

F Britto, Sophia University, Tokyo, Japan

L. Vishwanatha Ramaswami Iyer (also Aiyar) was born on October 25, 1895, in Thrissur, Kerala. He did his Intermediate (pre-university) studies as a State Scholar at Maharaja's College, Ernakulam. He received his B.A. (1914) in geology as a student of Presidency College, Chennai, and his B.L. (1916) in law, as a student of Law College, Chennai. After a brief period of legal practice and high school teaching, he earned an M.A. in English language and literature, and he joined his alma mater Maharaja's College on November 29, 1922, as an assistant professor of English. He spent the rest of his life at Maharaja's as a teacher and researcher, and died – aged only 52 – on January 31, 1948, in Thrissur, while on sick leave.

Known to his acquaintances as LVR, Iyer made lasting contributions to Dravidian Linguistics, even though he had no formal education in linguistics and his job required no mastery of Dravidian matters. With a flair for languages and a passion for linguistics, he learned as many as 18 languages (Indian and non-Indian, literary and non-literary), read avidly all the linguistics publications available to him, and refined his linguistics skills through personal contacts with several leading linguists, such as Suniti Kumar Chatterji. He preferred to write articles and monographs rather than lengthy books and published them in a wide variety of Indian and international journals. Limiting himself to Dravidian linguistics, he wrote – according to an estimate – nearly 200 articles on language evolution, word derivation, language reconstruction, phonetics, phonology, morphology, etymology, syntax, and semantics. His primary interests were Malayalam, Tamil, and Tulu, but he also wrote on and drew on many other languages such as Greek, Kannada, Telugu, Kui, Kuvi, Kurukh (Kurux), Gondi, and Brahui. He penned numerous book reviews and translated books and articles, including Romain Rolland's life of Gandhi from French to English.

The most noteworthy of Iyer's works are those he wrote on (1) *Lilatilakam*, a grammatical treatise of the 14th century, and its relation to Malayalam and Tamil; (2) the phonetics, phonology, morphology, and evolution of Malayalam; (3) the Dravidic Sandhi; (4) the language of ancient Tamil literature; (5) the Dravidian – especially Tamil – verbs; (6) the language features of Tulu, Kui, and Kurukh; and (7) comparative Dravidian and the evolution of Dravidian languages. His works are treasured even today for the abundant data they contain and for the creative ways in which Iyer manipulates them. Since no other linguist had explored Dravidian languages as thoroughly as he did, he was a pioneer in many areas of Indian linguistics, and he was acknowledged, even during his lifetime, as a leading authority on Indian philology. Linguists concerned with India, such as Murray B. Emeneau, Jules Bloch, and Thomas Burrow, have held Iyer's works in high esteem, and Kamil V. Zvelebil, another eminent Dravidolgist, says of Iyer's writings that they "contain rare insights, valuable data, and path-breaking ideas."

There exists no published compilation of all the articles that Iyer wrote, and his few books (e.g., Iyer, 1925, 1936, 1944) are out of print and hard to get. A volume edited by Subramoniam and Nair (1978) brings together the essays of five Dravidian linguists on the contributions and limitations of Iyer. It also has a large bibliography of Iyer compiled by Gopala Pillai.

Bibliography

(n. d.). *Collected papers of L. V. R.* (vols I and II). [Cyclostyled copies.] Annamalai Nagar: Annamalai University Department of Linguistics.

Iyer L V R (1925). *A brief account of Malayalam phonetics.* Calcutta: University of Calcutta.

Iyer L V R (1936). *The evolution of Malayalam morphology.* Thrissur: Ramavarma Research Institute.

Iyer L V R (1944). *Grammar in Lilatilakam.* Thrissur: The Saraswathy Printing and Publishing House.

Ramaswami Iyer L V (1997). *Dravidian encyclopaedia. (vol. III) (Language and literature).* Thiruvananthapuram: The International School of Dravidian Linguistics 694.

Subramoniam V I & Nair B G (eds.) (1978). *L. V. Ramaswamy Iyer and Seshagiri Prabhu.* Thiruvananthapuram: Dravidian Linguistics Association.

J

Jaberg, Karl (1877–1959)

S N Dworkin, University of Michigan, Ann Arbor, MI, USA

Trained in Zurich under Emile Freymond and at the *École des Hautes Études* in Paris where he studied with Jules Gilliéron, Karl Jaberg spent his entire university teaching career (1907–1945) at the University of Bern. Dialect geography, lexical history, and semantics constitute the main components of his research and were masterfully combined in many of his historically slanted lexical studies. His dissertation (1901) dealt with the evolution of pejorative meanings in French, and his *Habilationsschrift* (1906) examined verbal inflection in a Franco-Provençal dialect. Unlike his teacher Gilliéron, Jaberg's work involved diachronic and comparative analysis of the synchronic distribution of lexical items. He was codirector (with Jakob Jud) of the *Sprach und Sachatlas Italiens und der Südschweiz* and the author of many studies on individual lexical items written from the perspective of diachronic onomasiology, *Wörter und Sachen*, and cultural history. In addition, Jaberg wrote on the broader methodological issues of dialect geography (see especially, Jaberg 1908, 1936; Jaberg and Jud, 1928). Questions concerning the psychological side of language as well as its affective and expressive use as a communicative instrument and how these characteristics could affect the formal development of a word also attracted Jaberg's attention. Toward the end of his career, he became interested in the effects on lexical history of popular beliefs and superstitions, e.g., his study on the Romance terms for the birthmark (Jaberg, 1957), his only piece published in English (original German version in Jaberg, 1965: 282–322).

Selections of Jaberg's essays have been collected in Jaberg (1937, 1965). Complete bibliographies of his writings are available in the supplement to Heinimann (1958: 9–18) and in Jaberg (1965: 325–334). Detailed discussion of his life and work can be found in Heinimann (1958), Malkiel (1959), Terracini (1959), Boléo (1960), and Hilty (1997).

See also: Gilliéron, Jules (1854–1926); Jud, Jakob (1882–1952).

Bibliography

Boléo M & de Paiva (1960). '*A vida e a obra de Karl Jaberg*.' *Revista portuguesa de filologia 10,* 419–449.

Heinimann S (1958). 'Karl Jaberg, 1877–1958.' *Vox Romanica 17,* 1–8.

Hilty G (1997). 'Karl Jaberg (1877–1958).' In Wuest J (ed.) *Les linguistes suisses et la variation linguistique. Actes d'un colloque organisé à l'occasion du centenaire du Séminaire des langues romanes de l'Université de Zurich.* Romanica Helvetica 116. Basel/Tübingen: Francke. 115–124.

Jaberg K (1901). *Pejorative Bedeutungsentwicklung im Französischen.* Halle. [Also published in *Zeitschrift für romanische Philologie 25,* 561–601; *27,* 25–71; *29,* 25–71.]

Jaberg K (1906). *Über die assoziativen Erscheinungen in der Verbalflexion einer südostfranzösichen Dialektgruppe. Eine prinzipielle Untersuchung.* Aarau.

Jaberg K (1908). *Sprachgeographie. Beitrag zum Verständnis des Atlas linguistique de la France.* Aarau.

Jaberg K (1936). *Aspects géographiques du langage.* Paris: E. Droz.

Jaberg K (1937). *Sprachwissenschaftliche Forschungen und Erlebnisse.* Romanica Helvetica 6. Zurich/ Leipzig: Droz.

Jaberg K (1957). 'The birthmark in folk belief, language, literature, and fashion.' *Romance Philology 10,* 307–342.

Jaberg K (1965). *Sprachwissenschaftliche Forschungen und Erlebnisse. Neue Folge.* In Heinimann S (ed.) Romanica Helvetica 75. Bern: Francke Verlag.

Jaberg K & Jud J (1928). *Der Sprachatlas als Forschungsinstrument. Kritische Grundlagen und Einführung in den Sprach- und Sachatlas Italiens und der Südschweiz.* Halle: Niemeyer.

Malkiel Y (1959). 'Necrology: Karl Jaberg.' *Romance Philology 12,* 258–261.

Terracini B (1959). '*In memoria di Karl Jaberg*.' *Archivo glottologico italiano 44,* 84–100.

Jackendoff, Ray S. (b. 1945)

J-M Fortis, C.N.R.S.– University of Paris 7, Paris, France

Ray S. Jackendoff was born on January 23, 1945, in Chicago (Illinois, USA). As a student of Chomsky, his work on syntax adopts the generativist framework. However, his interest for semantics has set him at the junction of multiple theoretical trends (generative grammar, cognitive science and semantics, philosophy of mind). He is now a professor at Brandeis University, near Boston.

In his first book (1972), he argued against specifying semantic conditions on transformations or representing semantic information in deep structure. Jackendoff defends a strict separation of syntactic processes and of semantic rules. He ascribes to the semantic component the function of relating deep structure to a semantic decomposition of lexical items and clauses and of providing rules of interpretation for anaphora, pragmatic structure, and arguments in modal contexts.

His activity veered thereafter toward syntactic matters. The outcome was a theory of phrase structure (1977), the X-bar syntax, in which all lexical categories were described as instantiating a common hierarchical pattern with three levels of constituency. X-bar syntax has since been incorporated in government and binding theory.

The idea that semantic generalizations must be formulated independently from syntactic ones was pursued in his *Semantics and cognition*. In this book, propositional representations ('conceptual structures' or CS) built out of semantic primitives clearly form a kind of *lingua mentalis*. Semantics, it is claimed, must ultimately reduce to cognitive psychology. CS also function as a *sensus commmunis*. Thus, visual and linguistic informations are said to communicate through CS. Most importantly, CS representing spatial relations are regarded as providing the machinery with which other fields (such as time or possession) can be conceptually organized and rendered expressible. The theory of CS is considerably elaborated in Jackendoff's 1990 book, which dealt with a number of semantic issues such as the conceptual representation of thematic roles, the description of different semantic fields (motion and causation), and the composition of CS of lexical items with those of adjuncts.

Since the 80s, Jackendoff's activity aims toward a description of the functional architecture of the language faculty and its interaction with other cognitive processes. His approach has remained remarkably consistent: the language faculty is segmented into three rule systems (on the phonological, syntactic, and conceptual levels) that communicate, through interface rules, with each other and with other processing systems (such as the perceptual systems). Two topics have been particularly explored: the relation of processing levels to conscious awareness (1987) and the interface with visual representations (notably the cerebral and cognitive constraints on the linguistic expression of spatial relations).

See also: Chomsky, Noam (b. 1928); Cognitive Semantics; Consciousness, Thought and Language; Generative Grammar; Language of Thought; Language, Visual Cognition and Motor Action; Meaning: Cognitive Dependency of Lexical Meaning; Semantic Primitives; X-Bar Theory.

Bibliography

Jackendoff R S (1972). *Semantic interpretation in generative grammar.* Cambridge, MA: M.I.T. Press.

Jackendoff R S (1977). *X-bar syntax: a study of phrase structure.* Cambridge, MA: M.I.T. Press.

Jackendoff R S (1983). *Semantics and cognition.* Cambridge, MA: M.I.T. Press.

Jackendoff R S (1987). *Consciousness and the computational mind.* Cambridge, MA: M.I.T. Press.

Jackendoff R S (1990). *Semantics structures.* Cambridge, MA: M.I.T. Press.

Jackendoff R S (2002). *Foundations of language: brain, meaning, grammar, evolution.* Oxford: Oxford University Press.

Jackson, Kenneth Hurlstone (1909–1991)

G Mac Eoin, National University of Ireland, Galway, Ireland

Kenneth Jackson was born in Surrey on November 1, 1909 to a business family. Having been educated locally, he was elected a scholar of St John's College, Cambridge, from which he graduated in 1931 in Classics. However, he switched to Celtic in the new department of Anglo-Saxon, Norse, and Celtic and graduated in 1933 with the Allen Research Student Scholarship, which was awarded to him again the following year. He spent his first year as a Research Student at the University of Bangor, North Wales, where he attended the lectures of Ifor Williams and Thomas Parry. The second year of his Research Studentship was spent in Dublin, where he studied under Osborn Bergin. He spent a month each year in Wales and in the Blasket Islands off the coast of County Kerry, Ireland. In both of these places, he began his publishing career with 'Coracle Fishing Terms' from Wales and '*Dhá Scéal ón mBlascaod*' from Kerry, which were published in 1933. He transcribed stories from Peig Sayers, the well-known storyteller of the Blaskets, in the International Phonetic Alphabet. He published during those years two books, *Early Welsh gnomic poems* and *Studies in early Celtic nature poetry* (both in 1935). On returning to Cambridge, he was elected Fellow of St John's College and appointed Lecturer in Celtic. In 1939, he was appointed Lecturer and later Associate Professor and Professor of Celtic at Harvard. During the war, he acted as censor on behalf of the British and American authorities in Bermuda. In 1950, he was appointed Professor of Celtic in Edinburgh, where he was to spend the rest of his life. By 1953, he had completed the work that was to bring him fame, *Language and history in early Britain*. This was an examination of the changes that took place in the early languages of Britain, in Anglo-Saxon, Irish Gaelic (Irish), Latin, and Welsh, from the 5th to the 7th centuries, their historical significance, and how they could be dated. It is a monumental work of scholarship. Jackson's next great work was *A historical phonology of Breton* (1967), an attempt to establish "a framework in which all the chief historical problems are raised and discussed." Again, it was a book of immense learning, but it has not received the same attention as *Language and history*, probably because of the limited number of scholars who study Breton in English-speaking countries. In addition to these works, Jackson edited a number of Irish and Welsh texts, principally *The Gododdin* (1969), *The Gaelic notes in the Book of Deer* (1972), and *Aislinge Meic Con Glinne* (1990), this last text containing the notes on Middle Irish that Jackson contributed as a replacement for the historical grammar of Irish that he had projected. Add to these a lecture that has led to much controversy, *The oldest Irish tradition: A window on the iron age* (1964), in which he claimed that the earliest Irish tales provide us with a vivid picture of life in the Iron Age, a claim hotly disputed. He died on February 20, 1991.

See also: Bergin, Osborn Joseph (1873–1950).

Bibliography

Williams J E Caerwyn (1979–1980). 'Kenneth Hurlstone Jackson.' *Studia Celtica XIV/XV,* 1–11 (Includes a list of publications by Kenneth Hurlstone Jackson).

Williams J E Caerwyn (1991–1992). 'Necrologie: Kenneth Hurlstone Jackson: 1909–1992.' *Studia Celtica XXVI/XXVII,* 201–212.

Jacobsen, Lis (1882–1961)

F Gregersen, Frederiksberg, Denmark

Elisabeth (Lis) Jacobsen, born on September 1, 1882 in Copenhagen, was a Danish runologist, language historian, lexicographer, and entrepreneur. Lis Jacobsen was born into the prominent Danish family of Jewish origins, the Rubins. Her cousin was the celebrated professor of psychology Edgar Rubin. While she was a student he and she joined a discussion group that included, among others, Niels and Harald Bohr. She was first educated as a teacher, but having married the philologist Jacob Peter Jacobsen in 1903, she began to study Nordic philology at the University of Copenhagen with his encouragement and was awarded the gold medal for a prize essay in 1908. She took her magister degree in 1908. In 1910 she was awarded the degree of doctor of philosophy for her dissertation on the sources of the Danish

Jackson, Kenneth Hurlstone (1909–1991)

G Mac Eoin, National University of Ireland, Galway, Ireland

Kenneth Jackson was born in Surrey on November 1, 1909 to a business family. Having been educated locally, he was elected a scholar of St John's College, Cambridge, from which he graduated in 1931 in Classics. However, he switched to Celtic in the new department of Anglo-Saxon, Norse, and Celtic and graduated in 1933 with the Allen Research Student Scholarship, which was awarded to him again the following year. He spent his first year as a Research Student at the University of Bangor, North Wales, where he attended the lectures of Ifor Williams and Thomas Parry. The second year of his Research Studentship was spent in Dublin, where he studied under Osborn Bergin. He spent a month each year in Wales and in the Blasket Islands off the coast of County Kerry, Ireland. In both of these places, he began his publishing career with 'Coracle Fishing Terms' from Wales and '*Dhá Scéal ón mBlascaod*' from Kerry, which were published in 1933. He transcribed stories from Peig Sayers, the well-known storyteller of the Blaskets, in the International Phonetic Alphabet. He published during those years two books, *Early Welsh gnomic poems* and *Studies in early Celtic nature poetry* (both in 1935). On returning to Cambridge, he was elected Fellow of St John's College and appointed Lecturer in Celtic. In 1939, he was appointed Lecturer and later Associate Professor and Professor of Celtic at Harvard. During the war, he acted as censor on behalf of the British and American authorities in Bermuda. In 1950, he was appointed Professor of Celtic in Edinburgh, where he was to spend the rest of his life. By 1953, he had completed the work that was to bring him fame, *Language and history in early Britain*. This was an examination of the changes that took place in the early languages of Britain, in Anglo-Saxon, Irish Gaelic (Irish), Latin, and Welsh, from the 5th to the 7th centuries, their historical significance, and how they could be dated. It is a monumental work of scholarship. Jackson's next great work was *A historical phonology of Breton* (1967), an attempt to establish "a framework in which all the chief historical problems are raised and discussed." Again, it was a book of immense learning, but it has not received the same attention as *Language and history,* probably because of the limited number of scholars who study Breton in English-speaking countries. In addition to these works, Jackson edited a number of Irish and Welsh texts, principally *The Gododdin* (1969), *The Gaelic notes in the Book of Deer* (1972), and *Aislinge Meic Con Glinne* (1990), this last text containing the notes on Middle Irish that Jackson contributed as a replacement for the historical grammar of Irish that he had projected. Add to these a lecture that has led to much controversy, *The oldest Irish tradition: A window on the iron age* (1964), in which he claimed that the earliest Irish tales provide us with a vivid picture of life in the Iron Age, a claim hotly disputed. He died on February 20, 1991.

See also: Bergin, Osborn Joseph (1873–1950).

Bibliography

Williams J E Caerwyn (1979–1980). 'Kenneth Hurlstone Jackson.' *Studia Celtica XIV/XV,* 1–11 (Includes a list of publications by Kenneth Hurlstone Jackson).

Williams J E Caerwyn (1991–1992). 'Necrologie: Kenneth Hurlstone Jackson: 1909–1992.' *Studia Celtica XXVI/XXVII,* 201–212.

Jacobsen, Lis (1882–1961)

F Gregersen, Frederiksberg, Denmark

Elisabeth (Lis) Jacobsen, born on September 1, 1882 in Copenhagen, was a Danish runologist, language historian, lexicographer, and entrepreneur. Lis Jacobsen was born into the prominent Danish family of Jewish origins, the Rubins. Her cousin was the celebrated professor of psychology Edgar Rubin. While she was a student he and she joined a discussion group that included, among others, Niels and Harald Bohr. She was first educated as a teacher, but having married the philologist Jacob Peter Jacobsen in 1903, she began to study Nordic philology at the University of Copenhagen with his encouragement and was awarded the gold medal for a prize essay in 1908. She took her magister degree in 1908. In 1910 she was awarded the degree of doctor of philosophy for her dissertation on the sources of the Danish

national norm. Typically, she argued that the language of the secular administrative body of the royal chancellery was more important for the creation of a Danish orthographical norm than was the bible translation of Christian III (1550). The work was intended to be the first volume in a comprehensive history of the Danish language, but Jacobsen soon realized that the sources needed for such a project were not available yet. In 1911 she consequently founded the Danske Sprog-og Litteraturselskab (DSL) as a society for the edition of dictionaries and early Danish texts. The DSL soon became, and has since then remained, the primary purveyor of critical editions of central Danish works as well as dictionaries based on a thorough examination of the available sources. Jacobsen never attained the affiliation with the university for which she was obviously qualified, but probably served the interests of the national linguistic scene better by creating her own power base outside the traditional institutions.

It takes a rather rare combination of fund-raising talents, organizational powers, and philological knowhow to be a linguistic entrepreneur on the grand scale, but Jacobsen had it all. She was vital in securing the establishment and quick completion of the great standard Danish dictionary *Ordbog over det danske Sprog* (*ODS*), published in 28 volumes from 1915 to1956, being its administrator and organizer from the beginning. She continued this line with the much smaller popular version *Nudansk ordbog* (*NO*), published in two volumes in 1952 and 1953, with a number of later editions. As a runologist, she completed the two-volume edition of L. F. A. Wimmer's *Danske Runeindskrifter* (with E. Moltke, 1941–1942) and wrote a classic paper on the interpretations and misinterpretations of runic inscriptions, 'Wimmers landmandsstene' (1926).

During the German occupation of Denmark she had to escape to Sweden, where she lectured on Danish and paved the way for the postwar Nordic cooperation between the national language boards. Jacobsen's vision of a history of Danish was meanwhile accomplished by her colleague Peter Skautrup in his five volumes on the history of the Danish language, *Det danske Sprogs Historie* (1942–1968), which was deeply and avowedly inspired by her sociolinguistic viewpoint and was published by the DSL. Nordic cooperation was also the key to the grand *Kulturhistorisk leksikon for nordisk middelalder* (1956–1978), which was conceived and planned by Jacobsen.

Jacobsen was first and foremost an organizer and entrepreneur, but she was also a linguist and philologist with a conception of language that united historical and sociolinguistic perspectives while respecting the traditional role of dialects in the Danish philological milieu. She ranks among the best linguistic authors with her refined style and her superb talent for fruitful polemics. She died on June 18, 1961.

Bibliography

Brøndsted M (1982). *Lis Jacobsen på hundredårsdagen.* DSL præsentationshæfte nr.11. Copenhagen: C. A. Reitzel.

Glahder (red.): *Runer og Rids. Festskrift til Lis Jacobsen 29. January 1952.* Rosenkilde og Bagger 1952.

Jacobsen L (1910). *Studier til det danske Rigssprogs Historie.* Copenhagen: Gyldendal.

Jacobsen L (1912). *Kvinde og Mand. En Sprogstudie* fra dansk Middelalder. Copenhagen: Gyldendal.

Jacobsen, L (1926). 'Wimmers Landmandsstene', in Jacobsen (1927): 13–48.

Jacobsen L (1927). *Dansk Sprog: Kritik og Studier* (selected papers 1912–1927). Copenhagen: Gyldendal.

Pedersen I L: Lis Jacobsen, *Dansk kvindebiografisk leksikon*, http://www.kvinfo.dk/side/170/bio/681/query/Lis%20Jacobsen/ with references.

Jagić, Vatroslav (1838–1923)

M N Dedaić, West Haven, CT, USA

Vatroslav Jagić was a famous Croatian philologist and a leading voice in Slavic philology during the late 19th and early 20th centuries. Jagić was born in the northern Croatian town of Varaždin on July 6, 1838, to a shoemaker's family. The trajectory of his professional life began when he left his hometown for Vienna to study classical philology and Slavic languages with Franc Miklošič, a famous teacher whom Jagić would succeed 25 years later. After graduation, he taught for nine years at a Gymnasium (high school) in Zagreb, Croatia, and after receiving a doctorate in philosophy from Leipzig University in Germany, Jagić was appointed Professor of Comparative Linguistics in Odessa, Ukraine. Three years later, in 1874, Jagić moved to Berlin, where he taught Slavic philology and revitalized interest for this discipline

national norm. Typically, she argued that the language of the secular administrative body of the royal chancellery was more important for the creation of a Danish orthographical norm than was the bible translation of Christian III (1550). The work was intended to be the first volume in a comprehensive history of the Danish language, but Jacobsen soon realized that the sources needed for such a project were not available yet. In 1911 she consequently founded the Danske Sprog-og Litteraturselskab (DSL) as a society for the edition of dictionaries and early Danish texts. The DSL soon became, and has since then remained, the primary purveyor of critical editions of central Danish works as well as dictionaries based on a thorough examination of the available sources. Jacobsen never attained the affiliation with the university for which she was obviously qualified, but probably served the interests of the national linguistic scene better by creating her own power base outside the traditional institutions.

It takes a rather rare combination of fund-raising talents, organizational powers, and philological knowhow to be a linguistic entrepreneur on the grand scale, but Jacobsen had it all. She was vital in securing the establishment and quick completion of the great standard Danish dictionary *Ordbog over det danske Sprog* (*ODS*), published in 28 volumes from 1915 to1956, being its administrator and organizer from the beginning. She continued this line with the much smaller popular version *Nudansk ordbog* (*NO*), published in two volumes in 1952 and 1953, with a number of later editions. As a runologist, she completed the two-volume edition of L. F. A. Wimmer's *Danske Runeindskrifter* (with E. Moltke, 1941–1942) and wrote a classic paper on the interpretations and misinterpretations of runic inscriptions, 'Wimmers landmandsstene' (1926).

During the German occupation of Denmark she had to escape to Sweden, where she lectured on Danish and paved the way for the postwar Nordic cooperation between the national language boards. Jacobsen's vision of a history of Danish was meanwhile accomplished by her colleague Peter Skautrup in his five volumes on the history of the Danish language, *Det danske Sprogs Historie* (1942–1968), which was deeply and avowedly inspired by her sociolinguistic viewpoint and was published by the DSL. Nordic cooperation was also the key to the grand *Kulturhistorisk leksikon for nordisk middelalder* (1956–1978), which was conceived and planned by Jacobsen.

Jacobsen was first and foremost an organizer and entrepreneur, but she was also a linguist and philologist with a conception of language that united historical and sociolinguistic perspectives while respecting the traditional role of dialects in the Danish philological milieu. She ranks among the best linguistic authors with her refined style and her superb talent for fruitful polemics. She died on June 18, 1961.

Bibliography

Brøndsted M (1982). *Lis Jacobsen på hundredårsdagen.* DSL præsentationshæfte nr.11. Copenhagen: C. A. Reitzel.

Glahder (red.): *Runer og Rids. Festskrift til Lis Jacobsen 29. January 1952.* Rosenkilde og Bagger 1952.

Jacobsen L (1910). *Studier til det danske Rigssprogs Historie.* Copenhagen: Gyldendal.

Jacobsen L (1912). *Kvinde og Mand. En Sprogstudie* fra dansk Middelalder. Copenhagen: Gyldendal.

Jacobsen, L (1926). 'Wimmers Landmandsstene', in Jacobsen (1927): 13–48.

Jacobsen L (1927). *Dansk Sprog: Kritik og Studier* (selected papers 1912–1927). Copenhagen: Gyldendal.

Pedersen I L: Lis Jacobsen, *Dansk kvindebiografisk leksikon,* http://www.kvinfo.dk/side/170/bio/681/query/Lis%20Jacobsen/ with references.

Jagić, Vatroslav (1838–1923)

M N Dedaić, West Haven, CT, USA

Vatroslav Jagić was a famous Croatian philologist and a leading voice in Slavic philology during the late 19th and early 20th centuries. Jagić was born in the northern Croatian town of Varaždin on July 6, 1838, to a shoemaker's family. The trajectory of his professional life began when he left his hometown for Vienna to study classical philology and Slavic languages with Franc Miklošič, a famous teacher whom Jagić would succeed 25 years later. After graduation, he taught for nine years at a Gymnasium (high school) in Zagreb, Croatia, and after receiving a doctorate in philosophy from Leipzig University in Germany, Jagić was appointed Professor of Comparative Linguistics in Odessa, Ukraine. Three years later, in 1874, Jagić moved to Berlin, where he taught Slavic philology and revitalized interest for this discipline

among scholars. After six years in Berlin, Jagić spent another six years teaching at the university in Saint Petersburg, Russia, and in 1886 moved to Vienna as the successor to his teacher Miklošič. In that era, Vienna was the center of philology; Jagić's home, where he continued teaching Slavic seminars after retirement, was a gathering place for the most famous Central and East European Slavists of the time. In his later days, Jagić was appointed member of the Upper House of the Austrian Parliament.

Jagić's understanding of philology was manifested in the journal *Archiv für slavische Philologie* (*Archives for Slavic Philology*), which he established in 1875 with the intention to make it the principal forum for Slavic philology. Jagić's aim was to widen the borders of philology to include, besides language itself, linguistic and literary artifacts such as old scripts and monuments, as well as Slavic folklore and culture. Jagić was faithful to this goal during his 45-year editorial reign. At the time when Croatia was part of the Austro-Hungarian monarchy, Jagić embraced the political ideas and views of Bishop Josip Juraj Strossmayer, who believed that raising the level of a nation's culture is the only way to freedom. With Franjo Rački, Ivan Kukuljević, and others, Jagić worked on preserving Illyrian traditions under the South Slavic (Yugoslav) name. This orientation was exemplified by Jagić's interest in the life and work of Juraj Križanić (1618–1683), a Dominican priest who was known for his promotion of pan-Slavism and cooperation of the Catholic and Orthodox Churches. Jagić's most important publications (among about 600 attributed to him) are *Historija književnosti naroda hrvatskoga ili srpskoga* (*History of the literature of the Croatian or Serbian people), Tajni jezik u Slavena (Secret language among Slavs),* and *Prilozi slavenskoj sintaksi (Contributions to Slavic syntax)*. His studies of Old Church Slavonic argued that it did not originate in the central plains of Pannonia as most experts claimed, but rather in southern Macedonia. Jagić was also famous for his commentaries on mediaeval Slavic texts and correspondence of famous Slavic philologists (Dobrovski and Kopitar, among others). Jagić was influential as a literary critic, as he argued for realistic depiction of life in literature. He died on August 5, 1923, in Vienna, Austria.

See also: Miklošič, Franc (1813–1891); Old Church Slavonic; Serbian–Croatian–Bosnian Linguistic Complex; Slavic Languages.

Bibliography

Damjanović S (1988). *Opširnost bez površnosti: podsjetnik na život i djelo Vatroslava Jagića.* Zagreb: Međunarodni slavistički centar SR Hrvatske.

Frangeš I *et al.* (1986). *Jagićev zbornik.* Zagreb: Zavod za znanost o književnosti Filozofskog fakulteta.

Jagić V (1908). *Jagić-Festschrift: zbornik u slavu Vatroslava Jagića.* Berlin: Weidmannsche Buchhandlung.

Jagić V (1910). *Istoriia slavianskoi filologii. Trud ordinarnago akademika I. V. IAgicha.* Sanktpeterburg, Tip. I. Akademii nauk.

Jagić V (1930–1934). *Spomeni mojega života.* Posebna izdanja (Srpska akademija nauka i umetnosti), knj. 75, 104. Društveni i istorijski spisi; knj. 30, 45. Beograd: Srpska kraljevska akademija.

Jagić V (1963). *Pisma I. V. IAgicha k russkim uchenym 1865–1886.* Izd. podgotovili G P Blok i T I Lysenko. Pod red. V V Vinogradova i G P Bloka. Leningrad, Izd-vo Akademii nauk USSR [Leningradskoe otd-nie].

Katičić R (1991). 'Dalek put Vatroslava Jagića.' *Scientia 17(1–4),* 53–58.

Magner T F (1968). *Jagic and the Varazdin Kajkavian dialect.* New York: New York University Press/London: University of London Press.

Pasarić J (1899). *Jagić prema hrvatstvu i srpstvu.* Zagreb.

Plepelić Z (1995). 'Vatroslav Jagić und die Koenigliche Bibliothek zu Berlin.' *Smotra/Rundschau 1(2),* 111–114.

Salzmann-Ćelan M (2003). 'Jagić i Strossmayer: kodifikacija jugoslavenstva.' *Riječ 9(2),* 179–200.

Jakobson, Roman (1896–1982)

N Kerecuk, London, UK

Roman Jakobson (**Figure 1**), an American philologist and linguist of Russian Jewish descent, was born in Moscow. He moved to Prague in 1920 and to the United States during World War II. He graduated from Moscow University and participated in the Russian avant-garde movement in literature. When he moved to Prague, he pursued his doctorate at the University of Prague. From 1933 to 1939, he taught at the Massaryck University, Brno, but left Czechoslovakia during the Nazi occupation. He then traveled to Uppsala in Sweden, to Denmark, and to Norway before leaving for New York in 1941. He taught at the École Libre de Hautes Études, a French/Belgium free university accommodated by the New

among scholars. After six years in Berlin, Jagić spent another six years teaching at the university in Saint Petersburg, Russia, and in 1886 moved to Vienna as the successor to his teacher Miklošič. In that era, Vienna was the center of philology; Jagić's home, where he continued teaching Slavic seminars after retirement, was a gathering place for the most famous Central and East European Slavists of the time. In his later days, Jagić was appointed member of the Upper House of the Austrian Parliament.

Jagić's understanding of philology was manifested in the journal *Archiv für slavische Philologie* (*Archives for Slavic Philology*), which he established in 1875 with the intention to make it the principal forum for Slavic philology. Jagić's aim was to widen the borders of philology to include, besides language itself, linguistic and literary artifacts such as old scripts and monuments, as well as Slavic folklore and culture. Jagić was faithful to this goal during his 45-year editorial reign. At the time when Croatia was part of the Austro-Hungarian monarchy, Jagić embraced the political ideas and views of Bishop Josip Juraj Strossmayer, who believed that raising the level of a nation's culture is the only way to freedom. With Franjo Rački, Ivan Kukuljević, and others, Jagić worked on preserving Illyrian traditions under the South Slavic (Yugoslav) name. This orientation was exemplified by Jagić's interest in the life and work of Juraj Križanić (1618–1683), a Dominican priest who was known for his promotion of pan-Slavism and cooperation of the Catholic and Orthodox Churches. Jagić's most important publications (among about 600 attributed to him) are *Historija književnosti naroda hrvatskoga ili srpskoga* (*History of the literature of the Croatian or Serbian people), Tajni jezik u Slavena (Secret language among Slavs),* and *Prilozi slavenskoj sintaksi (Contributions to Slavic syntax)*. His studies of Old Church Slavonic argued that it did not originate in the central plains of Pannonia as most experts claimed, but rather in southern Macedonia. Jagić was also famous for his commentaries on mediaeval Slavic texts and correspondence of famous Slavic philologists (Dobrovski and Kopitar, among others). Jagić was influential as a literary critic, as he argued for realistic depiction of life in literature. He died on August 5, 1923, in Vienna, Austria.

See also: Miklošič, Franc (1813–1891); Old Church Slavonic; Serbian–Croatian–Bosnian Linguistic Complex; Slavic Languages.

Bibliography

Damjanović S (1988). *Opširnost bez površnosti: podsjetnik na život i djelo Vatroslava Jagića.* Zagreb: Međunarodni slavistički centar SR Hrvatske.

Frangeš I *et al.* (1986). *Jagićev zbornik*. Zagreb: Zavod za znanost o književnosti Filozofskog fakulteta.

Jagić V (1908). *Jagić-Festschrift: zbornik u slavu Vatroslava Jagića.* Berlin: Weidmannsche Buchhandlung.

Jagić V (1910). *Istoriia slavianskoi filologii. Trud ordinarnago akademika I. V. IAgicha.* Sanktpeterburg, Tip. I. Akademii nauk.

Jagić V (1930–1934). *Spomeni mojega života.* Posebna izdanja (Srpska akademija nauka i umetnosti), knj. 75, 104. Društveni i istorijski spisi; knj. 30, 45. Beograd: Srpska kraljevska akademija.

Jagić V (1963). *Pisma I. V. IAgicha k russkim uchenym 1865–1886.* Izd. podgotovili G P Blok i T I Lysenko. Pod red. V V Vinogradova i G P Bloka. Leningrad, Izd-vo Akademii nauk USSR [Leningradskoe otd-nie].

Katičić R (1991). 'Dalek put Vatroslava Jagića.' *Scientia 17(1–4),* 53–58.

Magner T F (1968). *Jagic and the Varazdin Kajkavian dialect.* New York: New York University Press/London: University of London Press.

Pasarić J (1899). *Jagić prema hrvatstvu i srpstvu.* Zagreb.

Plepelić Z (1995). 'Vatroslav Jagić und die Koenigliche Bibliothek zu Berlin.' *Smotra/Rundschau 1(2),* 111–114.

Salzmann-Ćelan M (2003). 'Jagić i Strossmayer: kodifikacija jugoslavenstva.' *Riječ 9(2),* 179–200.

Jakobson, Roman (1896–1982)

N Kerecuk, London, UK

Roman Jakobson (Figure 1), an American philologist and linguist of Russian Jewish descent, was born in Moscow. He moved to Prague in 1920 and to the United States during World War II. He graduated from Moscow University and participated in the Russian avant-garde movement in literature. When he moved to Prague, he pursued his doctorate at the University of Prague. From 1933 to 1939, he taught at the Massaryck University, Brno, but left Czechoslovakia during the Nazi occupation. He then traveled to Uppsala in Sweden, to Denmark, and to Norway before leaving for New York in 1941. He taught at the École Libre de Hautes Études, a French/Belgium free university accommodated by the New



Figure 1 Roman Jakobson (1896–1982).

School for Social Research in New York. He then taught at Columbia University (1943–1949) before becoming a professor at Harvard in 1949 and at MIT in 1957. In 1982, Jakobson died in Boston, Massachusetts.

He was a prolific author, and much of his work was dedicated to literary studies, but his role in the development of structural linguistics is also important. There have been many studies of his publications in English, in particular through the twentieth century, but significantly less of what he published in Russian has been examined. Most of these primary sources are now available for research.

In Moscow, Jakobson, along with the other young intellectuals, and G. O. Vinokur and P. Bogatyrev in particular, were involved in the creation of the Moscow Linguistic Circle, taking part in the debates that began to be developed in the second half of the nineteenth century on Slavic philology, poetics, dialectology, and folklore in the Old Russian Empire. After the death of F. E. Korsh (1843–1915), Jakobson chaired the Moscow Linguistic Circle until his departure for Prague. Traditionally, F. F. Fortunatov's Moscow School advocated a formal, logical approach to language study. A somewhat different treatment of the same themes predominated in St. Petersburg, where the OPOIAZ – Society for the Study of Poetic Language – was created in 1916 (it ceased to exist in 1923). Jakobson was a member of OPOIAZ, along with V. Shklosky, I. Tynianov, and B. Eikhenbaum. In St. Petersburg, the 'Kazan' School of J. Baudouin de Courtenay (1845–1929) and his disciples predominated. Both the Moscow Linguistic Circle and OPOIAZ are usually referred to as the Russian Formalists. However, influential studies in language, linguistics, and literature were not limited to these two centers in the Old Russian Empire and the Soviet Union: the city of Kharkiv, with its university, was fundamental, as was the legacy of O. O. Potebnia, with his disciples, in Moscow, St. Petersburg, and elsewhere, and the activities of the Historical-Philological Society (created in 1876).

After moving to Prague, Jakobson met Prince Nikolai Trubetzkoy (1890–1938) at a time when Prague was a hub of intellectual activity enabled by the intellectual climate fostered by T. G. Massaryk. Jakobson's main interests in the period continued to center on Slavic philology and literary studies. He was one of the founding members of the Prague Linguistic Circle, along with V. Mathesius (who was president until his death in 1945) and N. Trubetskoy, S. Karchevskii, and J. Murařosvký, among others. The Prague and Copenhagen centers became known as schools of structural linguistics. The former benefited from the Jakobson's move to the United States, where he promoted it.

The focus of Jakobson's writings during his Moscow and Prague periods is primarily aimed at literature, new art aesthetics, and philology (rather than linguistics), including Slavic philology, comparative philology, sound patterns and phonemics, literary language, the influence of folk literature on literature, folk songs, Futurism, aesthetics, poetics, versification, Russian and Slavonic culture and comparative mythology. It was after his arrival in the United States that Jakobson increased his output in phonology and, later, in linguistics, but he continued to describe himself as a Russian philologist.

There is still much research to be pursued on the evolution and sources of his ideas. For example, during the Moscow/Prague period, Jakobson wrote on literary themes and authors such as Trediakovsky (1915), Khlebnikov (1920, using his pen name R. A. [Roman Aljagrov] and for his own poetry), Maiakovsky (1931), Pushkin (1937), and Pasternak (1935).

Some of Jakobson's writings were dedicated to specific languages: Russian and its dialects, Czech, Bulgarian, and others. He also wrote on a number of linguists/philologists such as Durnovo (1926), Mathesius (1932), Shakhmatov (1920), and Peshkovskii (1933) (the latter two were disciples of O. O. Potebnia).

With P. Bogatyrev, he wrote on Slavic philology in Russia during the period of 1914–1921 (1922–1923). In the late 1920s, he produced two articles on phonetics and the phonological history of Slavonic languages. In 1931, he published a paper on the principles of historical phonology and entries on

'phoneme' and 'linguistics' for a Czech scientific dictionary in 1932 and 1935.

Jakobson was interested in identifying universals in languages, and he consequently disregarded relevant crucial differences. The idea of shared grammatical features as opposed to shared origin was part of his theory of universals. On the same basis, he postulated a development of the Slavonic verbal system (1932) based on a single stem. In 1936, he also published an article on the Russian/Slavonic case system based on the same idea. In 1938, he presented a paper on his theory of phonological affinities at the Fourth International Congress of Linguists in Copenhagen. In 1939, he wrote an article on the structure of phonemes. These pieces were the beginning of his work on phonology and linguistics.

In addition, in the 1920s and 1930s, Jakobson also had an interest in the sociology of language and national self-determination. These topics appeared in his papers in 1920–1921 and 1931. The latter year focuses on the proposed Eurasian linguistic union within the Eurasian ideological movement proposed by Russian émigrés in the 1920s and 1930s, characterized by an idiosyncratic linguistic and cultural approach.

His 1941 *Kindersprache, Aphasie und allgemeine Lautgesetze* (Jakobsen, 1971a: 328–401) has generated much research, as it suggests that there might be a universal order in the acquisition of speech sounds. His cooperation with N. Trubetskoi resulted in the furthering of the theory of distinctive features for the phonemes – the components of sounds. A phonetic feature distinguishes one phoneme from another (e.g., voiced/voiceless). The presupposition in this theory is that a set of features is applicable to distinguish phonological differences in all languages.

The concept of distinctive features, along with other concepts related to sound patterns, sound, and meaning, were expanded after Jakobson arrived in the United States. The Technical Report for the MIT Acoustics Laboratory, *Preliminaries to speech analysis* (1952) was jointly published with G. Fant and M. Halle. With Morris Halle, in 1956, in *Fundamentals of language*, the authors put forward the idea that a set of 12 binary oppositions, a bundle of 12 features, would be sufficient to give an account of all of the distinctions in all languages. The languages differ only in the manner that they combine these features. This contributed to the subsequent developments in phonological theory. For instance, A. Martinet outlined his views on Jakobson's phonology in his 'Substance phonique et traits distinctifs' (1957–1958).

Jakobson promoted the application of linguistics to literature. Some of the concepts that he used were those of iconicity, markedness, metaphor and metonymy, and communicative functions. He created a model of communicative functions based on the work of Karl Bühler and others. He argued that a shared code was not sufficient, but a context was essential in the communicative function. The components of communication are the addresser and the addressee and context, message, channel, and code. Therefore, depending on the focus of the component of communication, the functions are emotive (focuses on the addresser's attitude to his or her own message; e.g., interjections and emphatic speech), conative (focuses on the addressee; e.g., vocative), referential (refers to the context), phatic (refers to the contact/channel of communication between two speakers), metalinguistic (refers to the code itself, language about language; i.e., metalanguage), and poetic (refers to the additional component of a message apart from content). Each piece of discourse requires an analysis to identify which of the above functions predominate.

The seven volumes of Jakobson's *Selected writings* (1962–1984) offer an outline of the main areas of Jakobson's many interests. The correspondence of Jakobson with Trubetskoi and others in the interwar period (Toman, 1994) offers interesting insights into both the circulation of ideas and the characters of the personalities.

The acquaintance with C. Levi-Strauss has generated another strand of interaction and influence. The 1980 *Dialogues* with Krystyna Pomorska (his last wife) is another source. *The sound shape of language* with L. Waugh in 1986 represents another element in the history of Jakobson's thoughts. There are many publications on his life and works. Equally there is a great deal of comment about his works by Jakobson himself. Indeed, the accounts of intellectual histories in language sciences and the other areas that interested Jakobson are selective and have not been sufficiently researched. His eleven essays edited by K. Pomorska and S. Rudy (1985) *Verbal art, verbal sign, verbal time* are meant to be mainly an introduction to his works in poetics but also have relevance to his linguistic ideas.

See also: Bühler, Karl (1879–1963); Fant, (Carl) Gunnar (Michael) (b. 1919); Halle, Morris (b. 1923); Jakobson, Roman: Theory of the Sign; Levi-Strauss, Claude (b. 1908); Martinet, André (1908–1999); Mathesius, Vilém (1882–1945); Potebnja, Alexander (1835–1891); Shklovsky, Viktor Borisovich (1893–1984); Trubetskoy, Nikolai Sergeievich, Prince (1890–1938).

Bibliography

Aroux S (1996). *La philosophie du langage.* Paris: Presses Universitaires de France.

Culler J (1975). *Structuralist poetics. Structuralism, linguistics and the study of literature.* London: Routledge & Kegan Paul.
Davies A M (1998). *History of linguistics IV: Nineteenth-century linguistics.* London: Longman.
Iakobson R O (1996). 'Moskovskii lingvisticheskii kruzhok. Podgotovka teksta, publikatsiia, vstupetel'naia zametka i primechaniia M. I. Shapira.' *Philologica 3,* 361–379.
Jakobson R (1931). 'O fonologicheskikh iazykovykh soiuzakh.' In Jakobson R & Savytskii N S (eds.) *Evrasiia v Svete Iazykoznannia.* Paris. 7–12 [printed in Prague].
Jakobson R (1931). 'Über die phonologishchen Sprachbünde.' *Travaux de Cercle Linguistique de Prague 4,* 234–240.
Jakobson R (1971a). *Selected writings I. Phonological studies.* The Hague: Mouton de Gruyter.
Jakobson R (1971b). *Selected writings II. Word and language.* The Hague: Mouton de Gruyter.
Jakobson R (1981). *Selected writings III. Poetry of grammar and grammar of poetry.* The Hague: Mouton de Gruyter.
Jakobson R (1966). *Selected writings IV. Slavic epic studies.* The Hague: Mouton de Gruyter.
Jakobson R (1979). *Selected writings V. On verse. Its masters and explorers.* The Hague: Mouton de Gruyter.
Jakobson R (1984). *Selected writings VI. Early Slavic path and crossroads.* The Hague: Mouton de Gruyter.
Jakobson R (1985). *Selected writings VII. Contributions to comparative mythology. Studies in linguistics and philology.* The Hague: Mouton de Gruyter.
Jakobson R (1985). *Verbal art, verbal sign, verbal time.* Oxford: Basil Blackwell.
Jakobson R, Fant G M & Halle M (1952). *Preliminaries to speech analysis. The distinctive features and their correlates.* Cambridge, Mass.: MIT Press.
Kerecuk N (forthcoming). 'Annotated translation and bibliography of O. O. Potebnia's (1862).' *Thought and language.*
Koerner E F K (1999). *Linguistic historiography. Projects & prospects.* Amsterdam/Philadelphia: John Benjamins.
Koerner E F K (2000). *Historiographia linguistica 1973–1998. Indexes to volumes I–XXV.* Amsterdam/Philadelphia: John Benjamins.
Kristeva J [Julia Joayaux] (1969). *Le Langage, cet inconnu: une initiation à la linguistique.* Paris: Seuil.
Kukenheim L (1962). *Esquisse Historique de la linguistique française et de ses rapports avec la linguistique générale.* Leiden: Universitaire Pers.
Martinet A (1957–1958). 'Substance phonique et traits distinctifs.' In *Bulletin de la société linguistique de Paris.* 72–85.
Mattoso Camara J Jr. (1975). *História da lingüística.* Petrópolis: Editora Vozes.
Merleau-Ponty M (1964, 1973). *Consciousness and the acquisition of language.* Silverman H J (trans.). Evanston, IL: Northwestern University Press (1973). Originally published (1964) as 'La conscience et l'acquisition du langage.' *Bulletin de Psychologie 18(236),* 3–6.
Pomorska K *et al.* (1987). *Language, poetry and poetics. The generation of the 1890s: Jakobson, Trubetzkoy, Majakovskij.* Berlin: Mouton de Gruyter.
Robins R H (1967, 1997). *A short history of linguistics* (4th edn.). London: Longman.
Rudy S (1990). *Roman Jakobson 1896–1982: a complete bibliography of his writings.* Berlin: Mouton de Gruyter.
Stankiewicz E (1977). 'Roman Jakobson's work on the history of linguistics.' In Amstrong D & van Schooneveld C H (eds.) *Roman Jakobson: echoes of his scholarship.* Lisse/Holland: Peter de Ridder.
Steiner P (1984). *Russian formalism. A metapoetics.* Ithaca, NY: Cornell University Press.
Todorov T & Ducrot O (1972). *Dictionnaire encyclopèdique des sciences du language.* Paris: Seuil.
Toman J (1994). *Letters and other materials from the Moscow and Prague linguistic circles. 1912–1945 (=Cahiers Roman Jakobson, 1).* Ann Arbor: Michigan Slavic Publications.
Toman J (1995). *The magic of a common language. Jakobson, Mathesius, Troubetskoy, and the Prague linguistic circle.* Cambridge, MA: MIT Press.
Troubetskoy N S (1987). *Izbrannye trudy po filologii.* Moskva: Progress.

Jakobson, Roman: Theory of the Sign

M Danesi, University of Toronto, Toronto, Canada

Introduction

Roman Jakobson (1896–1982), the Moscow-born linguist and semiotician who carried out most of his work in the United States, is perhaps best known for his model of communication (Jakobson, 1960) – a model that is based on a social-dynamic theory of the sign. A student of Russian linguist Nikolai Trubetzkoy (1890–1939), Jakobson was a founding member of the Prague Linguistic Circle in 1920 (Koch, 1981: 225–226). Most of his writings on semiotics were published in the 1960s and 1970s, during his professorship at Harvard University (e.g., Jakobson, 1960, 1966, 1970, 1971, 1974, 1978, 1985).

Culler J (1975). *Structuralist poetics. Structuralism, linguistics and the study of literature.* London: Routledge & Kegan Paul.
Davies A M (1998). *History of linguistics IV: Nineteenth-century linguistics.* London: Longman.
Iakobson R O (1996). 'Moskovskii lingvisticheskii kruzhok. Podgotovka teksta, publikatsiia, vstupetel'naia zametka i primechaniia M. I. Shapira.' *Philologica 3,* 361–379.
Jakobson R (1931). 'O fonologicheskikh iazykovykh soiuzakh.' In Jakobson R & Savytskii N S (eds.) *Evrasiia v Svete Iazykoznannia.* Paris. 7–12 [printed in Prague].
Jakobson R (1931). 'Über die phonologishchen Sprachbünde.' *Travaux de Cercle Linguistique de Prague 4,* 234–240.
Jakobson R (1971a). *Selected writings I. Phonological studies.* The Hague: Mouton de Gruyter.
Jakobson R (1971b). *Selected writings II. Word and language.* The Hague: Mouton de Gruyter.
Jakobson R (1981). *Selected writings III. Poetry of grammar and grammar of poetry.* The Hague: Mouton de Gruyter.
Jakobson R (1966). *Selected writings IV. Slavic epic studies.* The Hague: Mouton de Gruyter.
Jakobson R (1979). *Selected writings V. On verse. Its masters and explorers.* The Hague: Mouton de Gruyter.
Jakobson R (1984). *Selected writings VI. Early Slavic path and crossroads.* The Hague: Mouton de Gruyter.
Jakobson R (1985). *Selected writings VII. Contributions to comparative mythology. Studies in linguistics and philology.* The Hague: Mouton de Gruyter.
Jakobson R (1985). *Verbal art, verbal sign, verbal time.* Oxford: Basil Blackwell.
Jakobson R, Fant G M & Halle M (1952). *Preliminaries to speech analysis. The distinctive features and their correlates.* Cambridge, Mass.: MIT Press.
Kerecuk N (forthcoming). 'Annotated translation and bibliography of O. O. Potebnia's (1862).' *Thought and language.*
Koerner E F K (1999). *Linguistic historiography. Projects & prospects.* Amsterdam/Philadelphia: John Benjamins.
Koerner E F K (2000). *Historiographia linguistica 1973–1998. Indexes to volumes I–XXV.* Amsterdam/Philadelphia: John Benjamins.
Kristeva J [Julia Joayaux] (1969). *Le Langage, cet inconnu: une initiation à la linguistique.* Paris: Seuil.
Kukenheim L (1962). *Esquisse Historique de la linguistique française et de ses rapports avec la linguistique générale.* Leiden: Universitaire Pers.
Martinet A (1957–1958). 'Substance phonique et traits distinctifs.' In *Bulletin de la société linguistique de Paris.* 72–85.
Mattoso Camara J Jr. (1975). *História da lingüística.* Petrópolis: Editora Vozes.
Merleau-Ponty M (1964, 1973). *Consciousness and the acquisition of language.* Silverman H J (trans.). Evanston, IL: Northwestern University Press (1973). Originally published (1964) as 'La conscience et l'acquisition du langage.' *Bulletin de Psychologie 18(236),* 3–6.
Pomorska K *et al.* (1987). *Language, poetry and poetics. The generation of the 1890s: Jakobson, Trubetzkoy, Majakovskij.* Berlin: Mouton de Gruyter.
Robins R H (1967, 1997). *A short history of linguistics* (4th edn.). London: Longman.
Rudy S (1990). *Roman Jakobson 1896–1982: a complete bibliography of his writings.* Berlin: Mouton de Gruyter.
Stankiewicz E (1977). 'Roman Jakobson's work on the history of linguistics.' In Amstrong D & van Schooneveld C H (eds.) *Roman Jakobson: echoes of his scholarship.* Lisse/Holland: Peter de Ridder.
Steiner P (1984). *Russian formalism. A metapoetics.* Ithaca, NY: Cornell University Press.
Todorov T & Ducrot O (1972). *Dictionnaire encyclopèdique des sciences du language.* Paris: Seuil.
Toman J (1994). *Letters and other materials from the Moscow and Prague linguistic circles. 1912–1945 (=Cahiers Roman Jakobson, 1).* Ann Arbor: Michigan Slavic Publications.
Toman J (1995). *The magic of a common language. Jakobson, Mathesius, Troubetskoy, and the Prague linguistic circle.* Cambridge, MA: MIT Press.
Troubetskoy N S (1987). *Izbrannye trudy po filologii.* Moskva: Progress.

Jakobson, Roman: Theory of the Sign

M Danesi, University of Toronto, Toronto, Canada

Introduction

Roman Jakobson (1896–1982), the Moscow-born linguist and semiotician who carried out most of his work in the United States, is perhaps best known for his model of communication (Jakobson, 1960) – a model that is based on a social-dynamic theory of the sign. A student of Russian linguist Nikolai Trubetzkoy (1890–1939), Jakobson was a founding member of the Prague Linguistic Circle in 1920 (Koch, 1981: 225–226). Most of his writings on semiotics were published in the 1960s and 1970s, during his professorship at Harvard University (e.g., Jakobson, 1960, 1966, 1970, 1971, 1974, 1978, 1985).

Theory of the Sign

Jakobson argued that the Saussurean view of the sign as an arbitrary structure – that is, one that has no motivation for being as it is through resemblance or simulation to its referent – was ultimately a useless one (*see* **Saussure: Theory of the Sign**). Taking his cue from Otto Jespersen (1860–1943), Jakobson saw signs as 'shifters' pointing to the cause and context of an utterance. Shifters are indicators of place (*here*, *there*), time (*now*, *then*), and specificity (*this*, *that*) that are context-sensitive.

The notion of shifter led to his model of communication, which has become widely used in both semiotics and linguistics. This consists of six constituents:

1. An addresser who initiates a communication;
2. A message that he or she wishes to convey and which he or she recognizes must refer to something other than itself;
3. An addressee who is the intended receiver of the message;
4. A context that provides the framework for encoding and decoding the message – e.g., the phrase 'Help me' takes on a different meaning depending on whether it is uttered by someone lying motionless on the ground or by someone in a classroom who is working on a difficult math problem;
5. A mode of contact (face-to-face, phone, etc.) by which the message is delivered between an addresser and an addressee;
6. A code (such as language, gesture, etc.) providing the signs for encoding and decoding messages.

Jakobson then pointed out that each of these constituents determines a different communicative function:

1. **Emotive**. This is the intent of the addresser in the communication act.
2. **Conative**. This is the effect the message has (or is intended to have) on its receiver.
3. **Referential**. This is any message that is constructed to convey information ('Main Street is two blocks north of here').
4. **Poetic**. This is any message constructed to deliver meanings effectively, like poetry ('I like Ike').
5. **Phatic**. This any message designed to establish, acknowledge, or reinforce social relations ('Hi, how's it going?').
6. **Metalingual**. This is any message that refers to the code used ('The word noun is a noun').

Conclusion

Jakobson's model suggests that sign use goes well beyond the function of simple information transfer. Signs are shifters that determine who says what to whom, where and when it is said, and how and why it is said; that is, sign use is motivated and shaped by the setting, the message contents, the participants, and the goals of each interlocutor. In-depth assessments of Jakobson's importance to semiotics can be found in Waugh (1976), Krampen (1981), and Eco (1985).

See also: Iconicity; Jakobson, Roman (1896–1982); Jespersen, Otto (1860–1943); Morris, Charles: Theory of the Sign; Saussure: Theory of the Sign; Semiotics: History; Structuralist Phonology: Prague School; Trubetskoy, Nikolai Sergeievich, Prince (1890–1938).

Bibliography

Eco U (1976). *A theory of semiotics.* Bloomington: Indiana University Press.

Eco U (1977). 'The influence of Roman Jakobson on the development of semiotics.' In Armstrong D & Schoonefeld C H van (eds.) *Roman Jakobson: echoes of his scholarship.* Lisse: The Peter de Ridder Press. 39–58.

Eco U (1979). *The role of the reader: explorations in the semiotics of texts.* Bloomington: Indiana University Press.

Eco U (1984). *Semiotics and the philosophy of language.* Bloomington: Indiana University Press.

Jakobson R (1960). 'Linguistics and poetics.' In Sebeok T A (ed.) *Style and language.* Cambridge, MA: MIT Press. 34–45.

Jakobson R (1966). 'Signe zero.' In Hamp E *et al.* (eds.) *Readings in linguistics II.* Chicago: University of Chicago Press. 109–115.

Jakobson R (1970). 'Language in relation to other communication systems.' In Olivetti C (ed.) *Linguaggi nella società e nella tecnica.* Milano: Edizioni di Comunità. 3–16.

Jakobson R (1971). *Selected writings II: Word and language.* The Hague: Mouton.

Jakobson R (1974). *Main trends in the science of language.* New York: Harper & Row.

Jakobson R (1978). *Six lectures on sound and meaning.* Mepham J (trans.). Cambridge, MA: MIT Press.

Jakobson R (1980). *The framework of language.* Ann Arbor: Michigan Studies in the Humanities.

Jakobson R (1985). *Selected writings VII.* Rudy S (ed.) Berlin: Mouton.

Koch W (1981). 'Roman Jakobson.' In Schnelle H (ed.) *Sprache und Gehirn.* Frankfurt: Suhrkamp. 223–236.

Krampen M (1981). 'A bouquet for Roman Jakobson.' *Semiotica 33*, 187–209.

Waugh L R (1976). *Roman Jakobson's science of language.* Lisse: Peter de Ridder.

Jakovlev, Nikolaj Feofanovič (1892–1974)

O Molchanova, Uniwersytet Szczecinski,
Szczecin, Poland

Nikolaj Feofanovič Jakovlev was born in the village of Bulgarin (in the Saratov region of the Russian Federation) on May 19, 1894, and died in Moscow on December 30, 1974. A Russian linguist, he was one of the founders of the Moscow School of Phonology and a specialist in Caucasian languages and the theory of orthography.

Jakovlev graduated from Moscow University in 1916. He became a professor in 1946 and obtained his Ph.D. in 1947. He held various posts at the Institute of Linguistics of the USSR Academy of Sciences in Moscow from researcher (1936) to Head of the Caucasian Languages Department at the same Institute (1942–1950). Jakovlev was also a member of the All-Union central committee of a new alphabet.

Phonology in Russia dated back to the 1870–1880s and was founded by Baudouen de Courtenay. However, its general theory was formulated by Ščerba, Jakovlev, Trubeckoj, and Jakobson half a century later.

Since 1920 Jakovlev has organized and led fieldwork trips to the Caucasus to study Caucasian languages that did not have written tradition and were used only for oral communication.

In 1923, on the basis of a linguistically consistent theory of a phoneme worked out by him, Jakovlev devised a mathematical formula for inventing systems of characters with which Caucasian languages could be written. It opened up new horizons for 70 Caucasian peoples for whom a new alphabet was first introduced on the basis of Latin and later Cyrillic.

He has also shown serious interest in fundamental linguistic problems of Caucasian languages concerning their grammatical structure and lexicon, which led Jakovlev to writing a grammar of the Kabardian-Cherkes language (1948). During his fieldwork in the Caucasus he gathered materials for the Kabardian language dictionary and got it published in 1927.

See also: Baudouin de Courtenay, Jan Ignacy Niecislaw (1845–1929); Jakobson, Roman (1896–1982).

Bibliography

Jakovlev N F (1925). *Inguši.* Moskva-Leningrad.
Jakovlev N F (1927). *Materialy dlja kabardinskogo slovarja.* Moskva.
Jakovlev N F (1931). 'Itogi unifikacii alfavitov v SSSR'. In *Sovetskoje stroitel'stvo № 8(16).*
Jakovlev N F (1940). *Sintaksis čečenskogo literaturnogo jazyka.* Moskva-Leningrad.
Jakovlev N F (1948). *Grammatika literaturnogo kabardinoerkesskogo jazyka.* Moskva-Leningrad.
Jakovlev N F (1970). 'Matematičeskaja formula postrojenija alfavita.' In Reformatskij A A (ed.) *Iz istorii otečestvennoj fonologii.* Moskva.

Jamaica: Language Situation

P L Patrick, University of Essex, Colchester, UK

Jamaica's 2.6 million people mostly speak the English-lexified Jamaican Creole, known locally as Patwa. (The Ethnologue name, Southwestern Caribbean Creole English, is used by neither Jamaicans nor creolists.) Patwa is a language of strong ethnic and national identification, spoken abroad in Canada, the UK, and the United States by well over 500 000 people. (On distinctive British varieties, see Sutcliffe, 1982; Sebba, 1993.) Patwa constitutes the basilect and mesolect of a Creole continuum (DeCamp, 1971); it is largely unintelligible to nonJamaicans. The widely-spoken acrolect, Standard Jamaican English, has emerging written/standard norms (Shields, 1989; Sand, 1999); though native to a small minority, education is largely in English. Patwa has increasingly penetrated broadcast and print media since independence in 1962. Jamaican Sign Language, apparently related to ASL but with some BSL-derived signs, is used in deaf schools.

Patwa arose from British colonialism, African slavery, and the plantation economy (Le Page, 1960). It developed between 1655, when English soldiers captured Jamaica, and 1750 – a century of massive increase in the African population – though some important features are not attested for another century (Lalla and D'Costa, 1990). Standard Jamaican is a recognized national variety of English, genetically descended by normal transmission from British dialects. Basilectal Patwa differs radically from English grammar, and emerged from profound language

Jakovlev, Nikolaj Feofanovič (1892–1974)

O Molchanova, Uniwersytet Szczecinski, Szczecin, Poland

Nikolaj Feofanovič Jakovlev was born in the village of Bulgarin (in the Saratov region of the Russian Federation) on May 19, 1894, and died in Moscow on December 30, 1974. A Russian linguist, he was one of the founders of the Moscow School of Phonology and a specialist in Caucasian languages and the theory of orthography.

Jakovlev graduated from Moscow University in 1916. He became a professor in 1946 and obtained his Ph.D. in 1947. He held various posts at the Institute of Linguistics of the USSR Academy of Sciences in Moscow from researcher (1936) to Head of the Caucasian Languages Department at the same Institute (1942–1950). Jakovlev was also a member of the All-Union central committee of a new alphabet.

Phonology in Russia dated back to the 1870–1880s and was founded by Baudouen de Courtenay. However, its general theory was formulated by Ščerba, Jakovlev, Trubeckoj, and Jakobson half a century later.

Since 1920 Jakovlev has organized and led fieldwork trips to the Caucasus to study Caucasian languages that did not have written tradition and were used only for oral communication.

In 1923, on the basis of a linguistically consistent theory of a phoneme worked out by him, Jakovlev devised a mathematical formula for inventing systems of characters with which Caucasian languages could be written. It opened up new horizons for 70 Caucasian peoples for whom a new alphabet was first introduced on the basis of Latin and later Cyrillic.

He has also shown serious interest in fundamental linguistic problems of Caucasian languages concerning their grammatical structure and lexicon, which led Jakovlev to writing a grammar of the Kabardian-Cherkes language (1948). During his fieldwork in the Caucasus he gathered materials for the Kabardian language dictionary and got it published in 1927.

See also: Baudouin de Courtenay, Jan Ignacy Niecislaw (1845–1929); Jakobson, Roman (1896–1982).

Bibliography

Jakovlev N F (1925). *Inguši.* Moskva-Leningrad.
Jakovlev N F (1927). *Materialy dlja kabardinskogo slovarja.* Moskva.
Jakovlev N F (1931). 'Itogi unifikacii alfavitov v SSSR'. In *Sovetskoje stroitel'stvo № 8(16).*
Jakovlev N F (1940). *Sintaksis čečenskogo literaturnogo jazyka.* Moskva-Leningrad.
Jakovlev N F (1948). *Grammatika literaturnogo kabardinoerkesskogo jazyka.* Moskva-Leningrad.
Jakovlev N F (1970). 'Matematičeskaja formula postrojenija alfavita.' In Reformatskij A A (ed.) *Iz istorii otečestvennoj fonologii.* Moskva.

Jamaica: Language Situation

P L Patrick, University of Essex, Colchester, UK

Jamaica's 2.6 million people mostly speak the English-lexified Jamaican Creole, known locally as Patwa. (The Ethnologue name, Southwestern Caribbean Creole English, is used by neither Jamaicans nor creolists.) Patwa is a language of strong ethnic and national identification, spoken abroad in Canada, the UK, and the United States by well over 500 000 people. (On distinctive British varieties, see Sutcliffe, 1982; Sebba, 1993.) Patwa constitutes the basilect and mesolect of a Creole continuum (DeCamp, 1971); it is largely unintelligible to nonJamaicans. The widely-spoken acrolect, Standard Jamaican English, has emerging written/standard norms (Shields, 1989; Sand, 1999); though native to a small minority, education is largely in English. Patwa has increasingly penetrated broadcast and print media since independence in 1962. Jamaican Sign Language, apparently related to ASL but with some BSL-derived signs, is used in deaf schools.

Patwa arose from British colonialism, African slavery, and the plantation economy (Le Page, 1960). It developed between 1655, when English soldiers captured Jamaica, and 1750 – a century of massive increase in the African population – though some important features are not attested for another century (Lalla and D'Costa, 1990). Standard Jamaican is a recognized national variety of English, genetically descended by normal transmission from British dialects. Basilectal Patwa differs radically from English grammar, and emerged from profound language

contact, structural mixing and creolization between Africans and English speakers (Alleyne, 1988); it is not genetically descended from English, nor from African languages, though they are historically linked.

Table 1 Implicational scale: variable Patwa features[a]

Creole 'C':	*/ **d** /*	*/ **t** /*	*pikni*	*no ben*	*nana*	*nyam*
Speaker 4	C	C	C	C	C	C
Speaker 3	C	C	C	C	C	*e*
Speaker 7	C	C	C	C	*e*	*e*
Speaker 2	C	C	C	*e*	*e*	*e*
Speaker 6	C	C	*e*	*e*	*e*	*e*
Speaker 1	C	*e*	*e*	*e*	*e*	*e*
Speaker 5	*e*	*e*	*e*	*e*	*e*	*e*
English *'e'*	/ ð /	/ θ /	*child*	*didn't*	*granny*	*eat*

Adapted from DeCamp, 1971: 355.
[a]This table illustrates unidimensional Creole continuum for Jamaican Patwa. It uses six surface forms (three lexical, two phonological, and one syntactic) which are unrelated to each other, and charts their distribution for seven Patwa speakers, in an implicational display (see Rickford, 2002 for discussion of this device). For each item, the Patwa element at the top of the column is matched with its English equivalent at the bottom. The table does not attempt to represent individual variability or intermediate (mesolectal) elements, but it does successfully stratify each speaker into a separate cell, contrasting the acrolect to the basilect, with a range of intermediate patterns.

Creolists disagree whether a pidgin developed in Jamaica, or existed previously on the African coast and was relexified; or whether Patwa developed by abrupt creolization. African sources included Kwa and Bantu languages, especially Akan Twi, partially maintained by Jamaican Maroons, rebel Africans who achieved independence from England in 1739. The ritual Maroon Spirit Language resembles Suriname Creoles (Bilby, 1983). British sources include non-standard regional English dialects from London, Bristol, West Midlands, Liverpool, Scots and Irish English. Late-arriving migrants (post-1845) of Indian, Chinese, and Syrian heritage neither influenced the language's genesis nor evolved ethnic dialects.

Patwa exhibits some regional variants, especially in western Jamaica; more salient are urban/rural and social variation (Patrick, 1999). Patwa was long subordinated to standard English by an elite minority controlling power and official culture. Limited access for most Jamaicans to standard English resulted in the (post-)Creole continuum, a chain of minimally distinct varieties stretching from the acrolect to the basilect (DeCamp, 1971). Discrete, categorical approaches (community bilingualism, standard/dialect, and diglossia) have proven inadequate to describe the Jamaican language situation (Table 1).

Demographically, the mesolect – the intermediate range of speech varieties – dominates. It incorporates

Table 2 Relative influence of constraints on past-marking, by speaker (Patrick, 1999: 265)[a]

	Did/Neva		*Verb inflection*			
	±*Punctual*	±*Anterior*	±*Punctual*	±*Anterior*	*Prec. Envt*	*Foll. Envt*
Roxy	...	...	∅	∅	❶	✸
Rose	...	...	∅	∅	❶	✸
Noel	...	...	∅	∅	❶	✸
Olive	...	...	∅	∅	❶	✸
Opal	...	...	∅	❶	∅	✸
Matty	✸	❶	∅	❶	∅	✸
Bigga	...	...	✸	∅	❶	≈
Tamas	✸	❶	✸	∅	❶	≈
Mina	✸	❶	✸	∅	❶	≈
Dinah	✸	❶	✸	∅	❶	≈

❶ Primary constraint
✸ Secondary constraint
≈ Equivocal constraint
∅ Not significant
...(Not applicable)
[a]Table 2 has similar goals as Table 1, but does take into account individual variation and mesolectal elements. It displays part of the variability in past-marking among twelve individual speakers from the city of Kingston, for both preverbal invariant Creole markers of past reference (mesolectal *did*, and its negative counterpart *neva*) and English-like verb inflection. (Not all speakers use both.) Instead of focusing on the occurrence of unrelated surface forms, however, it identifies the relative strength of linguistic constraints – both Creole- and English-valued, phonological and grammatical – governing elements of the same grammatical subsystem: past-marking in the urban Patwa mesolect. Though each speaker does not have a distinctive pattern, a range of patterns does occur, including (for some speakers) both English-like phonological forms and contexts, and classic Creole constraints (anteriority and punctuality, which affect both the use of invariant preverbal particles and verb inflection, sometimes for the same speaker). The table demonstrates a systematic but variable mesolectal grammer, and ordered inherent variation within the Patwa mesolect.

elements of English structure within a systematic but variable Creole grammar: it lacks categorical agreement, but possesses many English forms and inflections, governed by both English and Creole linguistic constraints (Table 2). It shows strong, stable sociolinguistic patterns of speech production and evaluation, with highly focused Patwa norms but diffuse English norms (Patrick, 1999). It is not produced synchronically through language mixing (although Jamaicans do codeswitch and styleshift), but predates the basilect.

Fundamental reference works include a dictionary (Cassidy and LePage, 1967) and transformational grammar (Bailey, 1966); contemporary descriptions include Devonish (2004) and Patrick (2004).

See also: Akan; Bantu Languages; Kwa Languages; Language Change and Language Contact; Pidgins and Creoles: Overview.

Bibliography

Alleyne Mervyn C (1988). *Roots of Jamaican culture.* London: Pluto Press.

Bailey B L (1966). *Jamaican Creole syntax*. Cambridge: Cambridge University Press.

Bilby K (1983). 'How the "older heads" talk: a Jamaican Maroon spirit possession language.' *Nieuwe West-Indische Gids 57(1–2)*, 37–88.

Cassidy F G & Le Page R B (eds.) (1967). *Dictionary of Jamaican English*. New York: Cambridge University Press.

DeCamp D (1971). 'Towards a generative analysis of a post-creole speech continuum.' In Hymes D (ed.) *Pidginization and creolization of languages*. Cambridge: Cambridge University Press. 349–370.

Devonish H (2004). 'Jamaican Creole: phonology.' In Kortman B, Schneider E W, Upton C, Mesthrie R & Burridge K (eds.) *A handbook of varieties of English, vol 1: phonology.* 2 vols. Berlin, New York: Mouton de Gruyter.

Lalla B & D'Costa J (1990). *Language in exile: three hundred years of Jamaican Creole.* Tuscaloosa: University of Alabama Press.

Le Page R B (1960). Creole Language Studies 1. *Jamaican Creole: an historical introduction.* London: Macmillan.

Patrick P L (1999). *Urban Jamaican Creole: variation in the mesolect.* Amsterdam, Philadelphia: John Benjamins.

Patrick P L (2004). 'Jamaican Creole: morphology and syntax.' In Kortman B, Schneider E W, Upton C, Mesthrie R & Burridge K (eds.) *A handbook of varieties of English, vol 2: Morphology and synta.* 2 vols. Berlin, New York: Mouton de Gruyter.

Patrick P L & Figueroa E (2002). 'Kiss-teeth.' *American Speech 77(4)*, 383–397.

Rickford J R (2002). 'Implicational scales.' In Chambers J K, Trudgill P & Schilling-Estes N (eds.) *The handbook of language variation and change.* Oxford: Blackwell. 142–167.

Sand A (1999). *Linguistic variation in Jamaica. A corpus-based study of radio and newspaper usage.* Tübingen: Narr.

Sebba M (1993). *London Jamaican: language systems in interaction.* London: Longman.

Shields K (1989). 'Standard English in Jamaica: a case of competing models.' *English World-Wide 10(1)*, 41–53.

Sutcliffe D (1982). *British Black English.* Oxford: Blackwell.

Japan: History of Linguistics

H Eto, Yokohama, Japan

Historically, language study in Japan has had two main streams: foreign-language study for the purpose of assimilating various new forms of knowledge from abroad, and inquiry into the mother tongue to protect national identity and manifest the excellence of Japanese culture. From this viewpoint, the history of linguistic thought in Japan is characterized by three major landmarks. First, in the Nara and Heian periods (710–1192), in the stream of Chinese influence on the Japanese language, awareness of the mother tongue arose with national identity. Second, in the Edo period (1603–1867), while *kangaku* (classical Chinese learning) and *yogaku* (Western learning) thrived, the *kokugaku* (national learning) movement gained ground, focusing on textual and exegetical study of Japanese classical literature based on language study. Third, in the course of Westernization after the Meiji restoration (1868), the *kokugogaku* (scientific study of Japanese) school assimilated the fruits of Western linguistics into Japanese study and explored their own methods.

This article lays emphasis on the tradition of inquiry into foreign languages and the mother tongue by Japanese intellectuals and scholars, weighing up the aforementioned three major periods in the history of language thought in Japan.

elements of English structure within a systematic but variable Creole grammar: it lacks categorical agreement, but possesses many English forms and inflections, governed by both English and Creole linguistic constraints (**Table 2**). It shows strong, stable sociolinguistic patterns of speech production and evaluation, with highly focused Patwa norms but diffuse English norms (Patrick, 1999). It is not produced synchronically through language mixing (although Jamaicans do codeswitch and styleshift), but predates the basilect.

Fundamental reference works include a dictionary (Cassidy and LePage, 1967) and transformational grammar (Bailey, 1966); contemporary descriptions include Devonish (2004) and Patrick (2004).

See also: Akan; Bantu Languages; Kwa Languages; Language Change and Language Contact; Pidgins and Creoles: Overview.

Bibliography

Alleyne Mervyn C (1988). *Roots of Jamaican culture.* London: Pluto Press.

Bailey B L (1966). *Jamaican Creole syntax.* Cambridge: Cambridge University Press.

Bilby K (1983). 'How the "older heads" talk: a Jamaican Maroon spirit possession language.' *Nieuwe West-Indische Gids 57(1–2),* 37–88.

Cassidy F G & Le Page R B (eds.) (1967). *Dictionary of Jamaican English.* New York: Cambridge University Press.

DeCamp D (1971). 'Towards a generative analysis of a post-creole speech continuum.' In Hymes D (ed.) *Pidginization and creolization of languages.* Cambridge: Cambridge University Press. 349–370.

Devonish H (2004). 'Jamaican Creole: phonology.' In Kortman B, Schneider E W, Upton C, Mesthrie R & Burridge K (eds.) *A handbook of varieties of English, vol 1: phonology.* 2 vols. Berlin, New York: Mouton de Gruyter.

Lalla B & D'Costa J (1990). *Language in exile: three hundred years of Jamaican Creole.* Tuscaloosa: University of Alabama Press.

Le Page R B (1960). Creole Language Studies 1. *Jamaican Creole: an historical introduction.* London: Macmillan.

Patrick P L (1999). *Urban Jamaican Creole: variation in the mesolect.* Amsterdam, Philadelphia: John Benjamins.

Patrick P L (2004). 'Jamaican Creole: morphology and syntax.' In Kortman B, Schneider E W, Upton C, Mesthrie R & Burridge K (eds.) *A handbook of varieties of English, vol 2: Morphology and synta.* 2 vols. Berlin, New York: Mouton de Gruyter.

Patrick P L & Figueroa E (2002). 'Kiss-teeth.' *American Speech 77(4),* 383–397.

Rickford J R (2002). 'Implicational scales.' In Chambers J K, Trudgill P & Schilling-Estes N (eds.) *The handbook of language variation and change.* Oxford: Blackwell. 142–167.

Sand A (1999). *Linguistic variation in Jamaica. A corpus-based study of radio and newspaper usage.* Tübingen: Narr.

Sebba M (1993). *London Jamaican: language systems in interaction.* London: Longman.

Shields K (1989). 'Standard English in Jamaica: a case of competing models.' *English World-Wide 10(1),* 41–53.

Sutcliffe D (1982). *British Black English.* Oxford: Blackwell.

Japan: History of Linguistics

H Eto, Yokohama, Japan

Historically, language study in Japan has had two main streams: foreign-language study for the purpose of assimilating various new forms of knowledge from abroad, and inquiry into the mother tongue to protect national identity and manifest the excellence of Japanese culture. From this viewpoint, the history of linguistic thought in Japan is characterized by three major landmarks. First, in the Nara and Heian periods (710–1192), in the stream of Chinese influence on the Japanese language, awareness of the mother tongue arose with national identity. Second, in the Edo period (1603–1867), while *kangaku* (classical Chinese learning) and *yogaku* (Western learning) thrived, the *kokugaku* (national learning) movement gained ground, focusing on textual and exegetical study of Japanese classical literature based on language study. Third, in the course of Westernization after the Meiji restoration (1868), the *kokugogaku* (scientific study of Japanese) school assimilated the fruits of Western linguistics into Japanese study and explored their own methods.

This article lays emphasis on the tradition of inquiry into foreign languages and the mother tongue by Japanese intellectuals and scholars, weighing up the aforementioned three major periods in the history of language thought in Japan.

Language Awareness and National Identity

In a narrow sense, there was no systematic inquiry into language in Japan before the *kokugaku* movement of the 18th century. In spite of Japan's long and rich literary tradition, expanding from the 8th century, no equivalent to Dionysius Thrax can be found in the early and medieval periods. However, the people had a keen awareness of the linguistic uniqueness of their mother tongue, which was closely related to a consciousness of national identity against highly advanced foreign civilizations.

Belief in *Kotodama*, or the Miraculous Power of Words

Ancient Japanese people recognized the miraculous power of words and worshipped *kotodama* (the spirit of words). In *Man'yoshu*, an anthology of poetry compiled in the late 8th century, poets admired Japan as the land where the soul of language flourishes, and the preface to *Kokinwakashu*, the first collection of Japanese verse, compiled by royal command in about 905, stated: "Japanese poetry can move heaven and earth, induce the gods and demons to appreciate the beauties of nature, render men and women harmonious, and soothe the hearts of doughty warriors." It was thus strongly believed that the sacred and miraculous power of the native Japanese language, if expressed in the form of a *waka* (a 31-syllable Japanese poem) or a *norito* (a Shinto ritual prayer), could exert a significant influence on people, the entire world, and even the gods.

Chinese Influence and Language Awareness

The Japanese writing system is indebted to Chinese, and no indigenous script has ever been developed. *Kojiki* (*Record of ancient matters*, 712) was transcribed in Chinese characters solely as phonetic signs, and *Nihonshoki* (*Chronicles of Japan*, 720) was written perfectly in classical Chinese style. It was not until the 10th century that the two syllabaries, *katakana* and *hiragana*, which were developed out of simplified Chinese characters, were widely used throughout the country.

It is not clear when *kanji* (Chinese characters) first reached Japan, but the scholar Wani from the Korean peninsula is said to have brought 11 volumes of Chinese writings, including the *Analects of Confucius*, as early as the 4th century. While the Japanese people of that time introduced overwhelmingly higher systematized philosophy and religious dogma such as Confucianism and Buddhism, they continued to respect their own traditional values. In terms of language consciousness, they continued worshipping the miraculous power of their mother tongue and used *kanji* for the 'male' domain of learning and writing, for example, official documents, scholarly texts, and formal letters, and *yamatokotoba* (the pure native language written in *hiragana*) for expressing things spiritual, emotional, sentimental, etc. Describing love and human nature, Lady Murasaki's 11th-century *Genji-monogatari* (*The tale of Genji*) was written almost entirely in *yamatokotoba*.

Language Study as National Learning

The dichotomy of using classical Chinese as the medium of official and scholarly writing and pure Japanese, untainted and unfertilized by Chinese, for emotional content remained for many centuries. From the 9th century onwards, therefore, language study was confined to classical Chinese texts. In addition, after the arrival of Buddhism, phonological study of *Bongo* (Sanskrit) in Buddhist scriptures was conducted exclusively by Buddhist monks. In either case, Japanese was not considered appropriate as a subject of rigorous study.

The first group with a particular interest in the Japanese language was the Portuguese missionaries, mostly Jesuits, who came to Japan in the 16th century. The Jesuit priests reported back to Europe on the excellence of Japanese people and culture and published books to encourage systematic study of the Japanese language. But the works were written neither in Japanese nor in Chinese, but in Portuguese, and had no intrinsic influence on language study by Japanese intellectuals.

Prelude to the *Kokugaku* Movement

In the 17th century *kangaku* reached its peak. In particular, the *kogaku* (study of antiquity) school, established and fostered by Ito Jinsai (1627–1705) and Ogyu Sorai (1666–1728), exceeded the level of classical scholarship in China at that time, emphasizing the original trust of Confucianism and rejecting the dogmatic interpretation of the Confucian canon by the Neo-Confucian tradition from Chu Hsi. Their approach to the Chinese texts was shaped by a philological and exegetical rigor to examine the original meaning of the texts by means of removing the arbitrary interpretations and commentaries that might obscure the original teachings of Confucius.

In addition to the flourishing *kangaku*, *yogaku*, represented by *rangaku* (Dutch learning), was triggered by the encounter with the Dutch, the only Westerners who were allowed to engage in foreign

trade under the national seclusion policy of the Tokugawa government. For this new learning in various fields such as medicine, chemistry, military science, etc., the *rangaku* scholars depended entirely on textual interpretation of Dutch books based on a direct translation method, which resulted in the learning of Dutch grammar growing in popularity. The translation in 1774 of *Kaitaishinsho* (*New book of anatomy*) by Sugita Genpaku (1733–1817) and others from a Dutch original was a remarkable achievement of the *rangaku* school.

The *Kokugaku* Movement and Its Characteristics

In the 18th century a fierce nativist reaction emerged against the dominance of foreign studies. An especially strong opposition came from the *kokugaku* movement against *kangaku*. Traditionally, four scholars form the lineage of the orthodox *kokugaku* school: Kada no Azumamaro (1669–1736), Kamo no Mabuchi (1697–1769), Motoori Norinaga (1730–1801), and Hirata Atsutane (1776–1843). Through the textual and exegetical study of Japanese classical literature and ancient writings, the *kokugaku* scholars insisted, for the purpose of identifying and appreciating typical Japanese mentality and morality, on a return to *yamatogokoro* (original Japanese ways of thinking), that is to a time before Japan had been corrupted by *karagokoro* (Chinese ways).

Among these *kokugaku* scholars, Norinaga's achievement deserves special attention. Influenced by a monk of the Shingon sect, Keichu (1640–1701), the forerunner of *kokugaku*, Norinaga conducted a philologically oriented investigation of ancient Japanese thought and culture to explain the peculiarity and preeminence of *yamatogokoro*. Through his study of *Man'yoshu* and *Genji-monogatari*, he re-evaluated the *mono-no-aware* (sensitivity in literary creation) of the ancient Japanese people to refute the claims of Confucian rationalism introduced at a later stage.

Norinaga's greatest linguistic achievement is the *Kojikiden* (1764–1798), in which he not only explicated the ancient Japanese system of morality, but also succeeded in showing how to read the text of *Kojiki*, written in Chinese script, as phonetic signs for *yamatokotoba*. By means of studying the ancient language and literature Norinaga contrived to grasp intuitively the world that the ancient people experienced, an attitude which parallels the approach of German philologists of the 19th century. Norinaga's inductive method in philological and exegetical studies of Japanese classical literature provided a model for many succeeding generations of scholars.

In the Wave of Modernization: From the Traditional Method to Scientific Study

During the period of modernization in the late 19th and early 20th centuries, Japan learned eagerly from the West the ideal forms for promoting science and technology. The leaders in various areas of that period imported from Western countries what they considered the best. Besides *rangaku*, therefore, many other European languages, particularly English, German (Standard German), and French, became popular for the practical reason of digesting the various new accomplishments of learning. In this intellectual climate, the new wave of philology and linguistics rushed into Japan from the West and extended its influence over the entire academic world. First, foreign scholars invited to the imperial universities initiated modern linguistics, and subsequently their students and the *ryugakusei* (scholars and students studying abroad to absorb the most advanced findings of science and technology) played an important role in the department of *hakugengaku* (philology and linguistics), with which linguistic sciences in Japan commenced.

Western Influence on Japanese Linguistics

The new generation of Japanese scholars devoted themselves to Western linguistics and the scientific study of the Japanese language. They were under the strong influence of German historical and comparative linguistics.

Bun'yu Nanjo (1849–1927) and Junjiro Takagushu (1866–1945), scholars of Sanskrit and Buddhism, studied as *ryugakusei* Sanskrit grammar under the guidance of Friedrich Max Müller (1823–1900) at Oxford. Though Sanskrit was known to Buddhist monks as early as the 9th century, they began new Buddhist studies directly from the Sanskrit texts based on the comparative method, not through the Chinese translation or assistance.

Kazutoshi Ueda (1867–1937), the first professor of linguistics at Tokyo Imperial University and considered the father of modern Japanese linguistics, studied the comparative and historical method under the Neogrammarians at Leipzig and thereby provided the foundation of the *kokugogaku* school in Japan. He initiated the historical study of Japanese phonology and lectured on the history of Japanese linguistics, which was crucial in establishing the popularity of German linguistics, particularly Hermann Paul's *Prinzipien*. On this basis, various branches of Japanese linguistics were established by Ueda and his students, e.g., Katsuji Fujioka (1872–1935) in Altaic linguistics, Izuru Shinmura (1876–1967) in lexicology, Shinkichi Hashimoto (1882–1945) in Japanese

grammar, and Kyosuke Kindaichi (1882–1971) in Ainu philology. The next generation of linguists, Hisanosuke Izui (1905–1983) and Harushige Kodzu (1908–1973), concentrated more on Indo-European and European linguistics. The Western influence on Japanese linguistics and general linguistics in Japan was so great that almost all the major linguistic works in European languages were translated into Japanese. Ferdinand de Saussure's *Cours de linguistique générale* (1916) was translated by Hideo Kobayashi (1903–1978) in 1928, the world's first translation of the work.

In the *kokugogaku* school, some scholars took a stand against the attempts to schematize Japanese linguistics according to the Western model. From a psychological standpoint, Motoki Tokieda (1900–1967) advocated the theory of *gengokateisetu* (theory of language as a process) in opposition to the structuralists' perspective, notably to Saussure's concept of language. In the *kokugaku* school, Shinobu Orikuchi (1887–1953), following Norinaga's approach to text interpretation, established a folkloric approach to the essence of the Japanese psyche.

After World War II, the linguistic tendency in Japan shifted from European influence to American influence. The younger scholars studied Bloomfieldian linguistics and later Chomskyan linguistics, and applied their accomplishments to Japanese language study. In this sense, present-day linguistic study in Japan proceeds in parallel with the world's dominant stream.

See also: Buddhism, Japanese; Fujioka, Katsuji (1872–1935); Japan: Language Situation; Japan: Writing System; Japanese; Kobayashi, Hideo (1903–1978); Kodzu, Harushige (1908–1973); Motoori Norinaga (1730–1801); Nanjio, Bunyiu (1849–1927); Tokieda, Motoki (1900–1967); Ueda, Kazutoshi (1867–1937).

Bibliography

Doi T (1976). *The study of language in Japan: a historical survey.* Tokyo: Shinozaki Shorin.

Izui H (1963). 'Recent trends in Japanese linguistics.' In Mohrmann C, Norman F & Sommerfelt A (eds.) *Trends in modern linguistics.* Utrecht: Spectrum. 38–58.

Kaiser S (1994). 'Japan: history of linguistic thought.' In Asher R E & Simpson J M Y (eds.) *The encyclopedia of language and linguistics 4.* Oxford: Pergamon. 1800–1804.

Kodzu H (1960). 'Linguistics in Japan, 1945–1958.' *Kratylos 5,* 1–9.

Sebeok T A (ed.) (1967). *Current trends in linguistics 2: linguistics in East Asia and South East Asia.* The Hague: Mouton.

Shimomiya T (1996). 'European tradition in the history of linguistics in Japan.' In Cram D, Linn A & Nowak E (eds.) *History of linguistics, 1996: selected papers from the Seventh International Conference on the History of the Language Sciences (ICHoLS VII), Oxford, 12–17 September, 1996, vol. 1: Traditions in linguistics worldwide.* Amsterdam: John Benjamins. 261–268.

Shimomiya T (1990). 'Indogermanistik in Japan.' *Kratylos 35,* 1–31.

Japan: Language Situation

S Iwasaki, University of California, Los Angeles, CA, USA

Japanese, as the official language of Japan, usually refers to the standard or common variety of the language. It was formed in the early 20th century, based on a variety of the Tokyo dialect. Establishing a standard language was essential at that time when Japan was modernizing itself after 250 years of isolationism. The standard language is the language of education and mass communication and is understood by over 126 million people throughout Japan.

The standard language is widespread, particularly in its written form. This does not mean, however, that the written convention has been completely formalized or stabilized. Since its establishment, the standard language has undergone various stages of modification, particularly in the area of *kanji* (Chinese characters) and *kana* (syllabary symbols) use. *Kana* symbols are phonetic in nature and are used together with *kanji* characters in sentences. However, until the end of World War II, these symbols represented the classical Japanese pronunciation. With the postwar reforms, they now reflect the sound of the modern language more accurately. A major problem in the writing system, however, has been the use of *kanji*, particularly its large inventory. Proposals were made to limit the number for practical purposes. The current 'List of Characters for General Use (*Toyo Kanji Hyo*)' recommends 1945 characters for daily use, though 3000 characters are used regularly in publications.

The use of honorific language and that of loanwords are also constant topics of debate for both

grammar, and Kyosuke Kindaichi (1882–1971) in Ainu philology. The next generation of linguists, Hisanosuke Izui (1905–1983) and Harushige Kodzu (1908–1973), concentrated more on Indo-European and European linguistics. The Western influence on Japanese linguistics and general linguistics in Japan was so great that almost all the major linguistic works in European languages were translated into Japanese. Ferdinand de Saussure's *Cours de linguistique générale* (1916) was translated by Hideo Kobayashi (1903–1978) in 1928, the world's first translation of the work.

In the *kokugogaku* school, some scholars took a stand against the attempts to schematize Japanese linguistics according to the Western model. From a psychological standpoint, Motoki Tokieda (1900–1967) advocated the theory of *gengokateisetu* (theory of language as a process) in opposition to the structuralists' perspective, notably to Saussure's concept of language. In the *kokugaku* school, Shinobu Orikuchi (1887–1953), following Norinaga's approach to text interpretation, established a folkloric approach to the essence of the Japanese psyche.

After World War II, the linguistic tendency in Japan shifted from European influence to American influence. The younger scholars studied Bloomfieldian linguistics and later Chomskyan linguistics, and applied their accomplishments to Japanese language study. In this sense, present-day linguistic study in Japan proceeds in parallel with the world's dominant stream.

See also: Buddhism, Japanese; Fujioka, Katsuji (1872–1935); Japan: Language Situation; Japan: Writing System; Japanese; Kobayashi, Hideo (1903–1978); Kodzu, Harushige (1908–1973); Motoori Norinaga (1730–1801); Nanjio, Bunyiu (1849–1927); Tokieda, Motoki (1900–1967); Ueda, Kazutoshi (1867–1937).

Bibliography

Doi T (1976). *The study of language in Japan: a historical survey.* Tokyo: Shinozaki Shorin.

Izui H (1963). 'Recent trends in Japanese linguistics.' In Mohrmann C, Norman F & Sommerfelt A (eds.) *Trends in modern linguistics.* Utrecht: Spectrum. 38–58.

Kaiser S (1994). 'Japan: history of linguistic thought.' In Asher R E & Simpson J M Y (eds.) *The encyclopedia of language and linguistics 4.* Oxford: Pergamon. 1800–1804.

Kodzu H (1960). 'Linguistics in Japan, 1945–1958.' *Kratylos 5,* 1–9.

Sebeok T A (ed.) (1967). *Current trends in linguistics 2: linguistics in East Asia and South East Asia.* The Hague: Mouton.

Shimomiya T (1996). 'European tradition in the history of linguistics in Japan.' In Cram D, Linn A & Nowak E (eds.) *History of linguistics, 1996: selected papers from the Seventh International Conference on the History of the Language Sciences (ICHoLS VII), Oxford, 12–17 September, 1996, vol. 1: Traditions in linguistics worldwide.* Amsterdam: John Benjamins. 261–268.

Shimomiya T (1990). 'Indogermanistik in Japan.' *Kratylos 35,* 1–31.

Japan: Language Situation

S Iwasaki, University of California, Los Angeles, CA, USA

Japanese, as the official language of Japan, usually refers to the standard or common variety of the language. It was formed in the early 20th century, based on a variety of the Tokyo dialect. Establishing a standard language was essential at that time when Japan was modernizing itself after 250 years of isolationism. The standard language is the language of education and mass communication and is understood by over 126 million people throughout Japan.

The standard language is widespread, particularly in its written form. This does not mean, however, that the written convention has been completely formalized or stabilized. Since its establishment, the standard language has undergone various stages of modification, particularly in the area of *kanji* (Chinese characters) and *kana* (syllabary symbols) use. *Kana* symbols are phonetic in nature and are used together with *kanji* characters in sentences. However, until the end of World War II, these symbols represented the classical Japanese pronunciation. With the postwar reforms, they now reflect the sound of the modern language more accurately. A major problem in the writing system, however, has been the use of *kanji*, particularly its large inventory. Proposals were made to limit the number for practical purposes. The current 'List of Characters for General Use (*Toyo Kanji Hyo*)' recommends 1945 characters for daily use, though 3000 characters are used regularly in publications.

The use of honorific language and that of loanwords are also constant topics of debate for both

written and spoken forms of the standard language. Modern Japanese inherited a complex honorific system developed fully in medieval Japan, which emphasized the difference in the hierarchical social status among the speaker/writer, listener/reader, and referents. Modern Japan, however, has moved toward a more egalitarian society, and the honorific system has been simplified. The system is now more concerned with the treatment of the addressee according to the speaker's relationship with the addressee and the level of formality of the speech situation.

English became an important foreign language in the Meiji Period (1868–1912). It has been the most influential supplier of loanwords in Modern Japanese, especially after World War II. The majority of the loanwords, which is about 10% of the vocabulary used in the language, are from English. Loanwords are usually pronounced according to the Japanese sound system, but some are pronounced in more innovative ways. Some loanwords are firmly established in the lexicon (*/rajio/* < 'radio,' */batto/* 'baseball bat'), while others are not (*/akusesu/* 'access,' */hippu hoppu/* 'hip hop'). While the use of loanwords is a solution to the accommodation of new ideas and products that are entering the life of Japanese people at an increasingly faster pace, the overuse of loanwords in the area of public services has become a social concern. Attempts have been made to replace some new loanwords with semantically more transparent Sino-Japanese based words.

Though many Japanese can and do speak the standard variety, some may also use their own regional dialects for daily life in their community. Dialect areas are first divided between the mainland and Ryukyu (Okinawa). The mainland dialect area is further divided into eastern, western, and southern dialects. The eastern dialects are spoken from the eastern to northeastern part of Honshu, the largest of the four major islands, and include the 12 million speakers of the Tokyo dialect. The western dialects are spoken in the west and southwest part of Honshu and on the smallest major island of Shikoku. They include two important subdialects: Kyoto dialect, a descendant of the language used in much of the classical literature from the 8th through 16th centuries, and Osaka dialect with approximately 8 million speakers. The southern dialects are spoken on Kyushu Island, the southern most major island.

Ryukyu dialects are spoken on the Okinawan archipelago extending off Kyushu Island to the east of Taiwan. Ryukyu dialects are completely unintelligible to the speakers of other dialects. There are 1.3 million inhabitants, but monolingual speakers of the dialects have almost completely disappeared; those over 50 years of age are proficient both in a vernacular and the standard Japanese, while the younger generation is rapidly moving toward a state of monolingualism in standard Japanese.

The dialect of Hokkaido, the northern-most major island, is a mixture of other dialects due to immigration from different parts of Japan since the mid-19th century driven by the Japanese government's development policy.

Ainu was once spoken on this island. It is unrelated to Japanese, though its genetic affiliation is not established. While the number of people of Ainu ethnicity is estimated at 24 000 or more, there are only few living native speakers of the language, partly due to the government's earlier assimilation policy, which included a ban on the use of Ainu. However, the Ainu New Law of 1997 formally recognized this ethnic group as an indigenous minority, and there is now a government-supported revival movement for the language and culture of Ainu.

Among the foreign languages spoken in Japan, Korean, Chinese, and English have special places. Many Koreans came to Japan during the period of Japan's annexation of Korea (ending in 1945) and remained in Japan afterward. Koreans are now the largest minority group with 1 million people, but the number of ethnic Koreans who can speak the language has been diminishing in spite of the Korean community's effort to teach students in bilingual schools at different levels of education.

Although Japanese learn a large number of Chinese characters and use Sino-Japanese vocabulary extensively in writing and speech, few can use Chinese as a spoken language. This is due to a long tradition that treats the Chinese language as a language of high culture, and not for oral communication. There are 50 000 ethnic Chinese consisting of an older generation of Cantonese speakers and a newer generation of Mandarin speakers. The younger ethnic Chinese do not speak Chinese as their first language, and some attend Chinese language school to maintain their linguistic heritage. The recent upsurge in the number of migrant workers, students and returnees from various parts of China, however, has changed the demographic and linguistic profile of Chinese in Japan.

English is not used as a general means for oral communication, and the proficiency in English is limited despite the fact that most Japanese have studied English for 6 years by the time they graduate from high school. There have been movements to change this trend.

See also: Ainu; Japanese.

Language Maps (Appendix 1): Map 103.

Bibliography

DeChicchis J (1995). 'The current state of the Ainu language.' In Maher J C & Yashiro K (eds.) *Multilingual Japan. Special issue of Journal of Multilingual and Multicultural Development 16*, 103–123.

Gottlieb N (2001). 'Language planning and policy in Japan.' In Gottlieb N & Chen P (eds.) *Language planning and language policy: East Asian perspectives.* Richmond: Cruzon. 21–48.

Gottlieb N & Chen P (eds.) (2001). *Language planning and language policy: East Asian perspectives.* Richmond: Curzon.

Honna N (1995). 'English in Japanese society: language within language.' In Maher J C & Yashiro K (eds.) *Multilingual Japan. Special Issue of Journal of Multilingual and Multicultural Development 16*, 45–62.

Inoue F (2001). *Nihongo wa ikinokoreru ka: Keizai gengogaku no shiten kara (Can the Japanese language survive? From the viewpoint of economic linguistics).* Tokyo: PHP.

Kaiser S (2003). 'Language and script in Japan and other East Asian countries: between insularity and technology.' In Maurais & Morris (eds.) *Languages in a globalising world.* Cambridge: Cambridge University Press. 188–202.

Maher J C (1995). 'The kakyo: Chinese in Japan.' In Maher J C & Yashiro K (eds.) *Multilingual Japan. Special isue of Journal of Multilingual and Multicultural Development 16*, 125–138.

Maher J C & Kawanishi Y (1995). 'On being there: Korean in Japan.' In Maher J C & Yashiro K (eds.) *Multilingual Japan. Special issue of Journal of Multilingual and Multicultural Development 16*, 87–101.

Maher J C & Macdonald G (eds.) (1995). *Diversity in Japanese culture and language.* London: Kegan Paul International.

Maher J C & Yashiro K (eds.) (1995). *Multilingual Japan. Special Issue of Journal of Multilingual and Multicultural Development* 16(1, 2).

Matsumori A (1995). 'Ryukyuan: past, present and future.' In Maher J C & Yashiro K (eds.) *Multilingual Japan. Special issue of Journal of Multilingual and Multicultural Development 16*, 19–44.

Tamura S (1999). *The Ainu Language.* Tokyo: Sanseido.

Twine N (1991). *Language and the modern state: the reform of written Japanese.* London: Routledge.

Japan: Writing System

J M Unger, The Ohio State University, Columbus, OH, USA

Description of the Current Writing System

The primary elements of modern Japanese writing are *kanji* 'Chinese characters' and two syllabaries called *hiragana* and *katakana* (collectively, *kana*). This mixture of character types is often described by saying that Japanese has three distinct scripts, but that is misleading. Any Japanese utterance can be written entirely in *hiragana* or *katakana*. Indeed, before 1945, either *hiragana* or *katakana* could be used as the default syllabary to supplement *kanji*. By contrast, not every Japanese utterance can be written with *kanji* alone. There once were forms of writing in which only *kanji* were used, but they are seldom encountered today. Moreover, modern spelling rules for *hiragana* and *katakana* differ, the two sets of *kana* are not freely interchangeable in ordinary usage, and roman letters and arabic numerals (A–Z, a–z, 0–9) as well as punctuation marks and typographical symbols of both Western and Japanese origin are now frequently used.

Kana

Despite the appellation 'syllabary,' the basic unit that *kana* represent is actually the mora. These are the so-called short vowels (V), with or without a preceding consonant (CV) or a consonant-glide combination (CyV); the mora nasal /N/ (a one-beat resonant with several conditioned allophones); and the mora obstruent /Q/, which precedes obstruents and causes a one-beat delay in their release. Japanese syllables consist of one or two (occasionally more) morae; the distinction between syllable and mora is essential for the description of pitch accent, present in most dialects but not indicated in standard orthography (for further details, see Vance, 1987).

Number and Form The two sets of *kana* contain 48 basic shapes, of which 47 are traditionally arrayed, as shown in **Table 1**. This arrangement, called the *gojūon-zu* or '50 sound chart' (despite its gaps), is first attested around 901. It is meant to be read from the upper right corner in Chinese style so that the order of morae is *a i u e o, ka ki ku ke ko, sa si su se so*, etc. This ordering was inspired by the *siddhaṃ* script for Sanskrit, and is currently the standard collating sequence for word sorting analogous to

Bibliography

DeChicchis J (1995). 'The current state of the Ainu language.' In Maher J C & Yashiro K (eds.) *Multilingual Japan. Special issue of Journal of Multilingual and Multicultural Development 16*, 103–123.

Gottlieb N (2001). 'Language planning and policy in Japan.' In Gottlieb N & Chen P (eds.) *Language planning and language policy: East Asian perspectives*. Richmond: Cruzon. 21–48.

Gottlieb N & Chen P (eds.) (2001). *Language planning and language policy: East Asian perspectives*. Richmond: Curzon.

Honna N (1995). 'English in Japanese society: language within language.' In Maher J C & Yashiro K (eds.) *Multilingual Japan. Special Issue of Journal of Multilingual and Multicultural Development 16*, 45–62.

Inoue F (2001). *Nihongo wa ikinokoreru ka: Keizai gengogaku no shiten kara (Can the Japanese language survive? From the viewpoint of economic linguistics)*. Tokyo: PHP.

Kaiser S (2003). 'Language and script in Japan and other East Asian countries: between insularity and technology.' In Maurais & Morris (eds.) *Languages in a globalising world*. Cambridge: Cambridge University Press. 188–202.

Maher J C (1995). 'The kakyo: Chinese in Japan.' In Maher J C & Yashiro K (eds.) *Multilingual Japan. Special isue of Journal of Multilingual and Multicultural Development 16*, 125–138.

Maher J C & Kawanishi Y (1995). 'On being there: Korean in Japan.' In Maher J C & Yashiro K (eds.) *Multilingual Japan. Special issue of Journal of Multilingual and Multicultural Development 16*, 87–101.

Maher J C & Macdonald G (eds.) (1995). *Diversity in Japanese culture and language*. London: Kegan Paul International.

Maher J C & Yashiro K (eds.) (1995). *Multilingual Japan. Special Issue of Journal of Multilingual and Multicultural Development* 16(1, 2).

Matsumori A (1995). 'Ryukyuan: past, present and future.' In Maher J C & Yashiro K (eds.) *Multilingual Japan. Special issue of Journal of Multilingual and Multicultural Development 16*, 19–44.

Tamura S (1999). *The Ainu Language*. Tokyo: Sanseido.

Twine N (1991). *Language and the modern state: the reform of written Japanese*. London: Routledge.

Japan: Writing System

J M Unger, The Ohio State University, Columbus, OH, USA

Description of the Current Writing System

The primary elements of modern Japanese writing are *kanji* 'Chinese characters' and two syllabaries called *hiragana* and *katakana* (collectively, *kana*). This mixture of character types is often described by saying that Japanese has three distinct scripts, but that is misleading. Any Japanese utterance can be written entirely in *hiragana* or *katakana*. Indeed, before 1945, either *hiragana* or *katakana* could be used as the default syllabary to supplement *kanji*. By contrast, not every Japanese utterance can be written with *kanji* alone. There once were forms of writing in which only *kanji* were used, but they are seldom encountered today. Moreover, modern spelling rules for *hiragana* and *katakana* differ, the two sets of *kana* are not freely interchangeable in ordinary usage, and roman letters and arabic numerals (A–Z, a–z, 0–9) as well as punctuation marks and typographical symbols of both Western and Japanese origin are now frequently used.

Kana

Despite the appellation 'syllabary,' the basic unit that *kana* represent is actually the mora. These are the so-called short vowels (V), with or without a preceding consonant (CV) or a consonant-glide combination (CyV); the mora nasal /N/ (a one-beat resonant with several conditioned allophones); and the mora obstruent /Q/, which precedes obstruents and causes a one-beat delay in their release. Japanese syllables consist of one or two (occasionally more) morae; the distinction between syllable and mora is essential for the description of pitch accent, present in most dialects but not indicated in standard orthography (for further details, see Vance, 1987).

Number and Form The two sets of *kana* contain 48 basic shapes, of which 47 are traditionally arrayed, as shown in Table 1. This arrangement, called the *gojūon-zu* or '50 sound chart' (despite its gaps), is first attested around 901. It is meant to be read from the upper right corner in Chinese style so that the order of morae is *a i u e o, ka ki ku ke ko, sa si su se so*, etc. This ordering was inspired by the *siddhaṃ* script for Sanskrit, and is currently the standard collating sequence for word sorting analogous to

Table 1 Modern *Kana* and their *Kanji* sources

w-		r-		y-		m-		h-		n-		t-		s-		k-		ø-		
わ	ワ	ら	ラ	や	ヤ	ま	マ	は	ハ	な	ナ	た	タ	さ	サ	か	カ	あ	ア	-a
和		良		也		末		波	八	奈		太	多	左	散	加		安	阿	
ゐ	ヰ	り	リ			み	ミ	ひ	ヒ	に	ニ	ち	チ	し	シ	き	キ	い	イ	-i
爲	井	利				美	三	比		仁		知	千	之		幾		以	伊	
		る	ル	ゆ	ユ	む	ム	ふ	フ	ぬ	ヌ	つ	ツ	す	ス	く	ク	う	ウ	-u
		留	流	由		武	牟	不		努	奴	川		寸	須	久		宇		
ゑ	ヱ	れ	レ			め	メ	へ	ヘ	ね	ネ	て	テ	せ	セ	け	ケ	え	エ	-e
惠	慧	礼				女		部		祢		天		世		計	介	衣	江	
を	ヲ	ろ	ロ	よ	ヨ	も	モ	ほ	ホ	の	ノ	と	ト	そ	ソ	こ	コ	お	オ	-o
遠	乎	呂		與		毛		保		乃		止		曾		己		於		

alphabetization. (Another collating sequence still occasionally encountered begins *i ro ha ni ho he to*, etc.)

In reading Table 1, note the following:

1. *Katakana* are shown to the right of the corresponding *hiragana* in the unshaded cells. The *kanji* from which the *kana* is/are believed to have arisen are shown underneath in shaded cells. Column headings give the initial consonant of the mora represented by the *kana* below (zero in the case of the first column); the row headings give the final vowel. N.B. /si/ = [ʃi], /ti/ = [tʃi], /tu/ = [tsɨ], /hi/ = [çi], /hu/ = [ɸɯ].
2. Morae of the form CyV are written by combining the appropriate Ci *kana* with a small version of the appropriate yV *kana*. For example, /kya/ きゃ キャ; /syu/ = [ʃɯ] しゅ シュ; /tyo/ = [tʃo̹] ちょ チョ. There once was a mora /kwa/, written くゎ and クヮ. This kind of digraphic notation has been extended to innovative morae such as /kwo/ クォ, /ye/ イェ, etc., which occur in new loanwords and so are seldom seen in *hiragana*.
3. A diacritic called *nigori* or *dakuten* indicates the voicing of an obstruent initial; e.g., /gi/ ぎ ギ; /zu/ ず ズ; /de/ で デ; and /bo/ ぼ ボ. Modern /h/ originated as /p/; hence the *kana* in the *h*-column are used for morae in *b*-. Modern /p/ is marked with another diacritic, called *maru* or *handakuten*; e.g., /pa/ ぱ パ.
4. The morae /wi we wo/ have merged with /i e o/. Only the *kana* for /wo/ is in current use (the *katakana* is seldom seen); see 'Historical Spellings' below.
5. The mora nasal /ɴ/ (usually romanized as syllable-final *n*) is written with *hiragana* ん and *katakana* ン (not part of the *gojūon-zu*). The mora obstruent /ǫ/ (usually romanized as a geminate) is notated with *hiragana* っ, *katakana* ッ – small versions of the *kana* for /tu/.

Function Since 1945, *katakana* have been used, like italics in English, for the transcription of unfamiliar loanwords and for emphasis; they were also used for writing telegram forms and are still seen in certain kinds of computer output. *Hiragana* are the default

kana for ordinary writing; they are used for postpositions, many adverbs, sometimes for nouns, and the variable endings of verbs and adjectives (in which case they are called *okurigana* 'send-off *kana*'). They are also the currently preferred *kana* set for notating the intended readings of *kanji* (*furigana*); e.g., the personal name 土井 can be Tsuchii or Doi, so it is sometimes necessary to write 土井(つちい) or 土井(どい) to make the reading clear. (Historically, *katakana* were adapted from *kanji* specifically to serve as *furigana*, but are thus used less often today.)

The modern notations for so-called long vowels and innovative morae in *hiragana* and *katakana* differ. For example, the syllable /too/ is written とう ⟨to u⟩ in *hiragana* (とお ⟨to o⟩ in a few words) but usually トー ⟨to ː⟩ in *katakana*, where the 'long' mark ー (*bō*) has no *hiragana* counterpart. Likewise, *hiragana* ねい ⟨ne i⟩ or, in a few words, ねえ ⟨ne e⟩ are both used for /nee/, whereas modern *katakana* usage prescribes ネー ⟨neː⟩. An example of an innovative mora is [ti]. In native and Sino-Japanese morphemes, /ti/ is realized [tʃi] and is represented by *hiragana* ち or *katakana* チ; to accomodate the [ti] some speakers produce in recent loanwords like [paːtiː] 'party' (cf. more conservative [paːtʃiː]), a new digraph notation is therefore used, viz. *katakana* ティ. (The parallel *hiragana* てぃ is possible but rare since innovative morae occur in loanwords.)

Kanji

In theory, any character used to write Chinese can be pressed into service in Japanese; but in practice, far fewer are actually seen. A handful known as *kokuji* 'national characters,' devised in Japan on the model of *kanji*, are usually lumped together with *kanji* despite their non-Chinese origin, e.g., 峠, which stands for the native word *tōge* 'mountain pass.'

Number and form The script reforms implemented in the period 1946–1959 included a governmentally authorized list of 1850 characters (the *tōyō kanji*) that capped the number to be used in official documents, educational curricula, and to a large extent in the public media. In 1980, the policy of limiting the number of *kanji* in general use was greatly weakened and a modified list of 1945 characters (the *jōyō kanji*) was issued. Despite this conservative backlash, however, *kanji* usage has been gradually declining for decades as the Japanese proclivity for borrowing new words has diluted its long-time cultural dependence on China. An analysis of 1 001 154 characters of running text in the 4245 Kyodo News Agency dispatches over 1 week in 1971 showed that only 46.1% of them were *kanji* even though those *kanji* accounted for 86.8% of the 2601 distinct character types in the same sample (Hayashi *et al.*, 1982: 206). Indeed, many *kanji* are seldom seen outside dictionaries: in 1963, the commonest 800 accounted for 90% of all those used in newspapers (Hayashi *et al.*, 1982: 245).

As part of the 1946–1959 reforms, the shapes of many of the *tōyō kanji* were simplified. Some changes were quite minor (adding or omitting a single stroke; connecting two disjoint strokes), but others were fairly drastic. For example, 沢 for 澤 (one reading of which is *sawa* 'mountain stream'), 昼 for 晝 (*hiru* 'noon'), 尽 for 盡 (*tsuku(su)* 'exert oneself, exhaust'). Although largely based on simplifications that had evolved in handwriting over the centuries, the Japanese changes are often different from the character simplifications later introduced in China. For instance, 傳 has been simplified to 伝 (*tsuta(eru)* 'tell, inform') in Japan but to 传 in the People's Republic of China.

Function Though the large number and complex shapes of the *kanji* used in Japanese are often remarked upon, the principal difficulty of the writing system lies in the fact that the typical *kanji* can take multiple 'readings.' Readings consist of one or more morae and generally fall into two categories: *kun* and *on*. *Kun* originated as glosses on *kanji*; they are Japanese words or parts of words that meant something like the literary Chinese words the *kanji* were used to write. *On* originated as Japanese attempts to pronounce Chinese syllables in different periods of history. The largest group, *kan-on*, were established in the 8th century; the less numerous *go-on*, accumulated in earlier centuries, now predominate in Buddhist terminology; a much smaller group, called *tōsō-on*, entered Japan together with Zen Buddhism in the 13th century; and there is a catch-all category called *kan'yō-on* 'customary *on*,' which lexicographers cannot otherwise explain. Today, *on* from different periods or *on* and *kun* may be found side by side within the reading of a single string of *kanji*.

The 1945 *jōyō kanji* can collectively take 2187 *on* and 1900 *kun*: 664 (34.1%) take just a single *on*; all the rest (apart from the few *kokuji* on the *jōyō kanji* list that take only one *kun*) have two or more readings (Hayashi *et al.*, 1982: 249–251). Moreover, the higher its frequency of occurrence, the more readings a *kanji* is likely to take, especially if allophonic variants are counted separately. For example, the heavily used *kanji* 日, usually read *hi* 'sun; day' in isolation, can also stand for the boldface portions of (in broad phonetics) [**ni**hon] or [**nip**pon] 'Japan,' [mai**nitʃi**] 'every day,' [**nik**ki] 'diary,' [**nitt**ʃuː] 'throughout the day,' [**niʃ**ʃoku] 'solar eclipse,' [**nis**seki] 'Japan Red Cross,' [sai**dʒitsu**] 'holiday,' [getsuyoː**bi**] 'Monday,' and [mik**ka**] '3 days; 3rd-day of a month.' In cases called *jukujikun* 'multiple character glosses,' two or

more *kanji* are glossed with a single native word that, at least synchronically, cannot be analyzed into a corresponding string of readings. This is the case, continuing the previous example, with 昨日 *kinō* 'yesterday,' 今日 *kyō* 'today,' 明日 *ashita* or *asu* 'tomorrow,' and such names as 春日 *Kasuga*, 日下 *Kusaka*, and 日向 *Hyūga*. The other side of this coin is the exploitation of an *on* or *kun* of a *kanji* to write part of a word with which it has no etymological connection. This kind of contrived usage (like English ⟨EZ⟩ for 'easy' or ⟨K9⟩ for 'canine') is called *ateji* and was especially popular in the Meiji period; e.g., *kurabu* 'club' (today just クラブ in *katakana*) was written 倶楽部 because the characters could be read individually as *ku*, *ra(ku)*, and *bu* and broadly hinted at the false etymology 'together + fun + place.'

A few *kanji* have readings that are, strictly speaking, neither *on* nor *kun*. The *kokuji* 働, which is 動 (an authentic *kanji*) with the addition of the so-called 'man' radical on the left, was invented to represent *hatara-* in the Japanese verb *hataraku* 'work'; later, with the pseudo-*on* *dō* (lifted from 動), it was used in the Sino-Japanese compound *rōdō* 労働 'labor'; cf. Chinese 勞動 (PRC simplified 劳动) *láodòng*. The readings *zero* 'zero' for 零, *pēji* 'page' for 頁, and *doru* 'dollar' for 弗 could likewise be called pseudo-*kun*.

Much psycholinguistic research has been devoted to the question of how readers apprehend and remember *kanji*. Although one continues to read that they function purely as logograms or even ideograms, the preponderance of evidence suggests otherwise (Matsunaga, 2002; Kess and Miyamoto, 1994, 1999; Paradis *et al*., 1985). The idea that *kanji* function logographically is made superficially plausible by the fact that a single native word, or a part thereof, can be used to gloss many different *kanji*. For instance, the verb *hakaru* はかる 'plan; measure; consult with' may also be written 図る, 計る, 測る, 量る, 諮る, or 謀る to emphasize the connotations 'design,' 'compute,' 'gauge,' 'size up,' 'confer,' or 'plot,' respectively, though in all likelihood, *hakaru* has always been a single polysemous verb, not a six-way homonym.

Moreover, what counts as the *kun* of a *kanji* is not always well defined. For instance, it is now common to write *wakaru* 'understand' 分かる, with 分 for *wa*, because, though the *-ka-* is invariant in all forms of this verb, it has a cognate *wakeru* 'divide, share.' But one still sometimes sees 分る, with 分 for *waka*. There are even some uninflected Sino-Japanese words that have become so nativized that they may be found entirely in *kana*; e.g., たいへん for 大変 *taihen* 'a lot.'

Punctuation, Alphanumerics, and Other Symbols

Japanese texts now regularly include a wide variety of symbols not covered by the rubrics of *kana* or *kanji*. Twine (1984) outlines the rise of punctuation in the Meiji period. For a description of the large number of symbols in the UNICODE standard counted as Japanese, see Lunde (1999).

Japanese texts are often formatted in traditional Chinese style. Vertical columns of characters run from the top of the page to the bottom, covering it from right to left. Page folds in this style are on the right side of the book, pamphlet, or newspaper. In much modern writing, the Western style has become more common, especially in academic work, where it frequently necessary to embed equations and other alphanumeric strings in Japanese-language text. Some Japanese symbols (e.g., *bō*) are rotated 90° counterclockwise to accommodate this format.

The most distinctive symbols are the Japanese comma 、and period 。, the word separator , *kana* repeat mark ゝ (ゞ when voicing is added), the *kanji* repeat symbol 々, and the paired quotation marks「 . . . 」and 『 . . . 』. (The Western pairing "..." has also been adapted in the form 〝...〟 in vertical format.) Especially in newspapers and magazines, abbreviations for units are often written as clusters of small *katakana*, e.g., kilo- ㌔, gram ㌘, meter ㍍, and percent ㌫. Besides the familiar (. . .) [. . .] { . . . }, many documents make graphic distinctions among a rich assortment of filled and outline bullets, and paired enclosure symbols, e.g.,【 . . . 】〔 . . . 〕〈 . . . 〉《 . . . 》. 'Right' and 'wrong' are conventionally ○ *maru* and × *peke* or *batsu*; with Chinese numerals, 〇 is used for 0. Also worth noting are the libel-dodging ○ ○ *maru-maru* 'so-and-so,' currency sign ¥ 'yen,' attention-grabbing ※ *komejirusi* 'rice sign,' and Post Office logo 〒 for 'postal code.' Encircled characters such ①②③, ㊤㊥㊦ 'top, middle, bottom' and ㊧㊨ 'left, right' are surprisingly common; icons like ☎ 'telephone number,' 📠 'fax number,' and 🖂 'mailing address' are gaining currency. The Japanese typographic imagination seems to know no bounds. (Many insights in Hansell's (2002) discussion of the integration of alphanumerics into Chinese writing carry over *mutatis mutandis* to Japanese.)

Alternative Methods

Two less common ways of writing Japanese should be mentioned if only because both belie the claim that the use of *kanji* and *kana* have rendered the Japanese insensitive to word boundaries. In both cases, the insertion of spaces between written words (*wakachigaki*) is an essential part of the system.

One is Japanese braille, which was developed in the 1880s and, unlike other braille systems, is based on *kana*. See Unger (1984) for further details, including a summary of braille *wakachigaki* rules.

The other is romanization, which now plays a more important part in Japanese writing than meets the

eye. Other than script reform advocates, few Japanese produce documents in romanized form, but because of word processors (*wāpuro*), a large and growing segment of the population types using a QWERTY-like keyboard. Although input systems are available that convert code numbers or keystroke sequences directly into individual *kanji*, *kana*, and other symbols, or allow the user to inscribe a character on a touch-sensitive tablet, the overwhelming majority of users input romanized Japanese and press function keys that instruct the software to convert these strings into the desired output characters. (Keyboards for *kana* typing are available but are used more by older, computer-shy novices.) Thus, from a psychological standpoint, most Japanese who use computers work in romanized Japanese even though they may never read extended romanized texts.

History of the Writing System

Apart from some artifacts produced in China that made their way to Japan in the early centuries of the first millennium, Chinese writing probably first came to Japan in the 5th century by way of visitors from the kingdom of Paekche, located in the southwest quadrant of the Korean peninsula. Attempts to explain the pictures found on bell-like bronze *dōtaku* of the Yayoi period (ca. 300 B.C.E.–300 C.E.) as primitive writing are unconvincing. The so-called *jindai moji* 'divine-age script' of Hirata Atsutane (1778–1843), sometimes still mentioned in crank literature, is obviously based on Korean *han'gŭl*, a product of the 15th century.

Origins of *Kana*

Kanji were used in many different ways in Old Japanese writing. The graphic deformation of those *kanji* used in the role now played by *kana* did not occur until the early 10th century. Before that time, *kana* were just ordinary *kanji* – now called *man'yōgana* after the 8th-century poetry anthology *Man'yōshū* – that were used to represent a mora (occasionally a string of two) because of a link between one of the readings of the *kanji* and the mora(e). In the 8th century, from which our oldest extant Japanese books date, several different *kanji* were used in the *kana* function for each syllable. Most of the *kanji* shown in Table 1 were typical *man'yōgana*. (*Hiragana* ん /N/ was derived from 无 after the *gojūon-zu* was already established. The source of *katakana* ン /N/ is uncertain, and sources other than those shown have been proposed for *katakana* ツ /tu/, サ /sa/, and *hiragana* つ /tu/.)

The number of the descendants of the different *man'yōgana* used for each syllable gradually dwindled over time, but as late as the 20th century, rival *kana* for the same syllable could be found in print. The now abandoned shapes are retroactively called *hentaigana* 'irregularly shaped *kana*.'

Historical Spellings (*Kanazukai*)

As might be expected, changes in *kana* spellings did not keep pace with changes in pronunciation over the centuries and became inconsistent as writers struggled to reconcile what they heard with the way they had learned to write. Until 1945, a system of rationalized etymological spellings, developed by late-18th/early-19th-century scholars, was in general use. The difference between this system and the one that has been in use (with a few minor changes) since the Occupation are shown in Table 2. Prewar Japanese were supposed to distinguish the *jō* of *jōdan* 'joke' じよう, *jōzu* 'skillful' じやう, *jōzetsu* 'verbose' ぜう, *jōbu* 'healthy' ぢやう, and *rokujō* '6 *tatami* [mats]' でふ in writing even though they all sounded alike. (Note also that the use of small versions of や *ya* ゆ *yu* よ *yo* did not become regularized until after the war.) Modern orthography has retained just a few anachronisms, most notably は *ha*, へ *he*, and を *wo* for the postpositions now pronounced *wa*, *e*, and *o*, respectively. Also, [dʒi] and [(d)zɨ] in the standard language are now written じ /zi/ and ず /zu/, respectively, in most contexts, although some of these syllables were once phonemically distinct /di/ and /du/ (still heard in some dialects), for which ぢ for /di/ and づ for /du/ are available.

Insofar as the modern spellings are much closer to present-day pronunciation, a great deal of Japanese phonological history can be surmised simply by comparing the old and new spellings. For example, original /p/ has lenited in initial position to /h/, with allophones [ç] before /i, y/, [ɸ] before /u/, and [h] elsewhere. In medial position, lenition caused /p/ to merge with original /w/; later, /w/ (from all sources) went to zero except before /a/. Most modern /p/ arose from a source other than original /p/; this is why /h/ rather than /p/ alternates with /b/ morphophonemically, just as /t/ alternates with /d/, /s/ with /z/, and /k/ with /g/. Of special importance is the analysis of *man'yōgana* usage, which shows that the language of the 8th-century capital area (Nara) had many more distinct syllables than later stages of the language (see Lange, 1973 for details).

Quasi-Chinese Methods

There are two obsolete methods of writing Japanese in which *kana* were used purely for annotational purposes or not at all. One is called *kanbun* 'Chinese text'. In principle, any literary (classical) Chinese text

Table 2 Pre-1945 *Kana* spellings

/ /	*Current*	*Obsolete*	/ /	*Current*	*Obsolete*	/ /	*Current*	*Obsolete*
i	い	ゐ ひ						
u	う	ふ						
e	え (へ)	ゑ へ						
o	お (を)	を ほ ふ						
ka	か	くわ	ga	が	ぐわ			
zi	じ	ぢ						
zu	ず	づ						
wa	わ (は)	は						
yū	ゆう	ゆふ いう いふ						
ō	おう	おほ わう あう はう						
kō	こう	こふ かう かふ くわう	gō	ごう	ごふ がう がふ ぐわう			
sō	そう	さう さふ	zō	ぞう	ざう ざふ			
tō	とう	たう たふ	dō	どう	だう だふ			
nō	のう	のふ なう なふ						
hō	ほう	ほふ はう はふ	bō	ぼう	ぼふ ばう ばふ	pō	ぽう	ぽふ ぱう ぱふ
mō	もう	まう						
yō	よう	やう えう えふ						
rō	ろう	ろふ らう らふ						
kyū	きゅう	きう きふ	gyū	ぎゅう	ぎう			
syū	しゅう	しう しふ	zyū	じゅう	じう じふ ぢゅう			
tyū	ちゅう	ちう						
nyū	にゅう	にう にふ						
hyū	ひゅう	ひう	byū	びゅう	びう			
ryū	りゅう	りう りふ						
kyō	きょう	きやう けう けふ	gyō	ぎょう	ぎやう げう げふ			
syō	しょう	しやう せう せふ	zyō	じょう	じやう ぜう ぢやう でう でふ			
tyō	ちょう	ちやう てう てふ						
nyō	にょう	ねう						
hyō	ひょう	ひやう へう	byō	びょう	びやう べう	pyō	ぴょう	ぴやう ぺう
myō	みょう	めう						
ryō	りょう	りやう れう れふ						

Table 3 Principal differences among Romanization systems

Kana (Katakana)	*Hepburn*	*Kunrei-shiki*	*Nippon-shiki*
クヮ グヮ	ka, ga (*older* kwa, gwa)	ka, ga	kwa, gwa
シ	shi	si	si
シャ シュ ショ	sha, shu, sho	sya, syu, syo	sya, syu, syo
ジャ ジュ ジョ	ja, ju, jo	zya, zyu, zyo	zya, zyu, zyo
チ	chi	ti	ti
チャ チュ チョ	cha, chu, cho	tya, tyu, tyo	tya, tyu, tyo
ヂャ ヂュ ヂョ	ja, ju, jo	zya, zyu, zyo	dya, dyu, dyo
ツ	tsu	tu	tu
ヅ	zu (*older* dzu)	zu	du
フ	fu	hu	hu
ヲ	o (*older* wo)	o	wo
ン	n (*older* m before b, m, or p)	n	n
トウ トオ トー	tōō, tōō, tōō	tô, tô, tô	tô, tô, tô
ネイ ネエ ネー	nei, nē, nē	nei, nê, nê	nei, nê, nê
シリーズ	shiriizu	sirîzu	sirîzu

could be read in a variety of literary Japanese according to rules for glossing individual *kanji* (occasionally strings of *kanji*), supplying postpositions and grammatical endings as needed, and permuting the resulting phrases in conformance with Japanese syntax. In practice, although some Japanese writers understood how to write proper literary Chinese, aberrant varieties arose in Japan. *Kana* notations and a system of special punctuation marks came to be used to guide the reader in deciphering *kanbun*. Crawcour (1965) remains the best introduction in English.

Sōrōbun refers to an epistolary variety of *kanbun*, used from the 15th through 19th centuries, characterized by the heavy use of the polite auxiliary verb *sōrō* < *samurau* 'serve' (cf. *samurai* 'one who serves'). Kiyose (1983) is an easily accessible introduction to *sōrōbun*.

Romanization

The first romanization system for Japanese was created in the 16th century by Portuguese missionaries. It affords us an unusually clear view of Late Middle Japanese phonology. (There are also some contemporary *han'gŭl* transcriptions of Japanese that are helpful in this regard.)

Before the 1880s, romanization of Japanese words in the West was not standardized. Around 1885, two well-thought-out systems were proposed: the one used in the third edition of James Curtis Hepburn's Japanese–English dictionary, and the Nippon-siki system of Tanakadate Aikitu. The government proclaimed a compromise between these two systems by Cabinet Order (*kunrei*) in the 1930s, reconfirmed in the 1950s. Today, ISO Japanese romanization reflects this *kunrei-siki* standard; journalists, politicians, and academics in the West tend to cling to the Hepburn system despite the use of *kunrei-siki* among geographers, scientists, and (with some variations) linguists.

Table 3 shows the differences among the three principal systems. The three rows at the bottom illustrate the treatment of so-called long vowels within single morphemes. Adjacent vowels at morpheme boundaries are properly written out mora by mora: e.g., *suu* 'suck' ≠ *sū* or *sû* 'number'; *tou* 'ask' ≠ *tō* or *tô* 'east.' In Hepburn, *ii* is preferred even when there is no morpheme boundary between the two morae (e.g., the personal name Iida), but *ī* is sometimes seen anyway (e.g., for '[the letter name] E'). In all three systems, an apostrophe follows *n* for the mora nasal /N/ (*hiragana* ん) before *y* or a vowel.

See also: China: Writing System; Japanese.

Bibliography

Chamberlain B H (2000 [1905]). *Collected works of Basil Hall Chamberlain. Major works 3: a practical introduction to the study of Japanese writing (moji no shirube).* Bristol: Ganesha/Tokyo: Edition Synapse.

Crawcour E S (1965). *An introduction to kambun.* Ann Arbor: Center for Japanese Studies, University of Michigan Press.

Erbaugh M S (ed.) (2002). *Difficult characters: interdisciplinary studies of Chinese and Japanese writing.* Columbus: National East Asian Language Resource Center.

Gottlieb N [Twine] (2000). *Word-processing technology in Japan:* kanji *and the keyboard.* Richmond: Curzon.

Hansell M (2002). 'Functional answers to structural problems in thinking about writing.' In Erbaugh (ed.). 124–176.

Hayashi Ō *et al.* (eds.) (1982). *Zusetsu Nihongo: gurafu de miru kotoba no sugata.* Tokyo: Kadokawa shoten. [Japanese explained in charts: the structure of language as seen in graphs.]

Kess J F & Miyamoto T (comps.) (1994). *Japanese psycholinguistics: a classified and annotated research bibliography.* Amsterdam, Philadelphia: Benjamins.

Kess J F & Miyamoto T (1999). *The Japanese mental lexicon: psycholinguistics studies of kana and kanji processing.* Philadelphia: John Benjamins.

Kida J (1994). *Nihongo daihakubutsukan: akuma no moji to tatakatta hitobito* [The great Japanese language museum: people who battled demon characters]. Tokyo: Jasuto Shisutemu [Just Systems].

Kiyose G N (1983). 'Sōrōbun.' In Reischauer E O *et al.* (eds.) *Kodansha Encyclopedia of Japan,* 9 vols. Tokyo: Kodansha. 7, 230–231.

Kokugo Gakkai (ed.) (1976). *Kokugo-shi shiryō-shū* [collection of materials on the history of the Japanese language]. Tokyo: Musashino shoin.

Kornicki P (1998). *The book in Japan: a cultural history from the beginnings to the nineteenth century.* Leiden, Boston: Brill.

Lange R A (1973). *The phonology of eighth-century Japanese: a reconstruction based upon written records.* Tokyo: Sophia University Press.

Lunde K (1999). *CJKV information processing.* Sebastopol, CA: O' Reilly & Associates.

Matsunaga S (1995). *The role of phonological coding in reading kanji: a research report and some pedagogical implications.* Technical Report 6. Honolulu: University of Hawai'i Second Language Teaching and Curriculum Center.

Matsunaga S (2002). 'Are Chinese characters ideographs? An argument from the psycholinguistic perspective.' In Erbaugh (ed.). 75–91.

Paradis M, Hagiwara H & Hildebrandt N (1985). *Neurolinguistic aspects of the Japanese writing system.* Orlando, FL: Academic Press.

Seeley C (1991). *A history of writing in Japan.* Leiden: Brill.

Tsukishima H (1981). *Nihongo no sekai 5: Kana.* [The world of Japanese 5: *kana*.] Tokyo: Chūō kōron-sha.

Twine [Gottlieb] N (1984). 'The adoption of punctuation in Japanese script.' *Visible Language 18(3),* 229–237.

Twine [Gottlieb] N (1991). *Language and the modern state: the reform of written Japanese.* London: Routledge.

Unger J M (1980). 'Etymology of the word /kana/.' *Papers in Japanese Linguistics 7,* 173–184.

Unger J M (1984). 'Japanese braille.' *Visible Language 18(3),* 254–266.

Unger J M (1987). *The fifth generation fallacy: why Japan is betting its future on artificial intelligence.* New York: Oxford University Press.

Unger J M (1996). *Literacy and script reform in Occupation Japan: reading between the lines.* New York: Oxford University Press.

Vance T J (1987). *An introduction to Japanese phonology.* Albany: State University of New York Press.

Yamada H (1983). 'Certain problems associated with the design of input keyboards for Japanese writing.' In Cooper W E (ed.) *Cognitive aspects of skilled typing.* New York: Springer. 305–407.

Japanese

M Shibatani, Rice University, Houston, TX, USA

Japanese is spoken by virtually the entire population of Japan – some 128 million people in 2004. In terms of native speakers, the number easily surpasses the number of speakers of the major European languages (German and French), ranking sixth among the languages of the world, after Chinese, English, Russian, Hindi, and Spanish. Yet, among the major languages of the world, Japanese occupies a unique position in a number of respects. Unlike the languages spoken on the European, American, and Asian continents, the Japanese language, being spoken in an island nation, is physically isolated from other languages. Also, unlike major European languages such as English and Spanish, Japanese is primarily spoken within the confines of its national boundaries, with no other country using it as either a first or second language. Moreover, Japanese is the only major world language for which the genetic affiliation to other languages and language families has not been conclusively proved.

As in the case of other language isolates, including Basque and Burushaski (and the geographical neighbors of Japanese, Ainu and Korean) (*see* **Ainu** and **Korean**), the genealogy of Japanese has been a perennial problem that has attracted the attention of both specialists and laymen alike. Hypotheses have been presented assigning Japanese to virtually all major language families. Although attempts to relate Japanese to the Altaic family have been most systematic and perhaps most persuasive (*see* **Altaic Languages**) two views have attracted increasing attention in recent years: (1) that Japanese consists of an Austronesian substratum and an Altaic superstratum, and (2) that Japanese is an Austronesian–Altaic hybrid or mixed language, in which not only simple lexical mixtures but also morphological hybrids, e.g., Austronesian verb roots with Altaic inflectional endings, are recognized (*see also* **Austronesian Languages: Overview**).

Kess J F & Miyamoto T (comps.) (1994). *Japanese psycholinguistics: a classified and annotated research bibliography.* Amsterdam, Philadelphia: Benjamins.
Kess J F & Miyamoto T (1999). *The Japanese mental lexicon: psycholinguistics studies of kana and kanji processing.* Philadelphia: John Benjamins.
Kida J (1994). *Nihongo daihakubutsukan: akuma no moji to tatakatta hitobito* [The great Japanese language museum: people who battled demon characters]. Tokyo: Jasuto Shisutemu [Just Systems].
Kiyose G N (1983). 'Sōrōbun.' In Reischauer E O *et al.* (eds.) *Kodansha Encyclopedia of Japan,* 9 vols. Tokyo: Kodansha. 7, 230–231.
Kokugo Gakkai (ed.) (1976). *Kokugo-shi shiryō-shū* [collection of materials on the history of the Japanese language]. Tokyo: Musashino shoin.
Kornicki P (1998). *The book in Japan: a cultural history from the beginnings to the nineteenth century.* Leiden, Boston: Brill.
Lange R A (1973). *The phonology of eighth-century Japanese: a reconstruction based upon written records.* Tokyo: Sophia University Press.
Lunde K (1999). *CJKV information processing.* Sebastopol, CA: O' Reilly & Associates.
Matsunaga S (1995). *The role of phonological coding in reading kanji: a research report and some pedagogical implications.* Technical Report 6. Honolulu: University of Hawai'i Second Language Teaching and Curriculum Center.
Matsunaga S (2002). 'Are Chinese characters ideographs? An argument from the psycholinguistic perspective.' In Erbaugh (ed.). 75–91.
Paradis M, Hagiwara H & Hildebrandt N (1985). *Neurolinguistic aspects of the Japanese writing system.* Orlando, FL: Academic Press.
Seeley C (1991). *A history of writing in Japan.* Leiden: Brill.
Tsukishima H (1981). *Nihongo no sekai 5: Kana.* [The world of Japanese 5: *kana*.] Tokyo: Chūō kōron-sha.
Twine [Gottlieb] N (1984). 'The adoption of punctuation in Japanese script.' *Visible Language 18(3),* 229–237.
Twine [Gottlieb] N (1991). *Language and the modern state: the reform of written Japanese.* London: Routledge.
Unger J M (1980). 'Etymology of the word /kana/.' *Papers in Japanese Linguistics 7,* 173–184.
Unger J M (1984). 'Japanese braille.' *Visible Language 18(3),* 254–266.
Unger J M (1987). *The fifth generation fallacy: why Japan is betting its future on artificial intelligence.* New York: Oxford University Press.
Unger J M (1996). *Literacy and script reform in Occupation Japan: reading between the lines.* New York: Oxford University Press.
Vance T J (1987). *An introduction to Japanese phonology.* Albany: State University of New York Press.
Yamada H (1983). 'Certain problems associated with the design of input keyboards for Japanese writing.' In Cooper W E (ed.) *Cognitive aspects of skilled typing.* New York: Springer. 305–407.

Japanese

M Shibatani, Rice University, Houston, TX, USA

Japanese is spoken by virtually the entire population of Japan – some 128 million people in 2004. In terms of native speakers, the number easily surpasses the number of speakers of the major European languages (German and French), ranking sixth among the languages of the world, after Chinese, English, Russian, Hindi, and Spanish. Yet, among the major languages of the world, Japanese occupies a unique position in a number of respects. Unlike the languages spoken on the European, American, and Asian continents, the Japanese language, being spoken in an island nation, is physically isolated from other languages. Also, unlike major European languages such as English and Spanish, Japanese is primarily spoken within the confines of its national boundaries, with no other country using it as either a first or second language. Moreover, Japanese is the only major world language for which the genetic affiliation to other languages and language families has not been conclusively proved.

As in the case of other language isolates, including Basque and Burushaski (and the geographical neighbors of Japanese, Ainu and Korean) (*see* **Ainu** and **Korean**), the genealogy of Japanese has been a perennial problem that has attracted the attention of both specialists and laymen alike. Hypotheses have been presented assigning Japanese to virtually all major language families. Although attempts to relate Japanese to the Altaic family have been most systematic and perhaps most persuasive (*see* **Altaic Languages**) two views have attracted increasing attention in recent years: (1) that Japanese consists of an Austronesian substratum and an Altaic superstratum, and (2) that Japanese is an Austronesian–Altaic hybrid or mixed language, in which not only simple lexical mixtures but also morphological hybrids, e.g., Austronesian verb roots with Altaic inflectional endings, are recognized (*see also* **Austronesian Languages: Overview**).

Among individual languages, Ryukyuan, Ainu, and Korean are the strongest candidates proposed as possible sister languages; in fact, the Japanese–Ryukyuan connection has been firmly established, and Ryukyuan, spoken in Okinawa, is now considered to be a dialect of Japanese (*see* **Ryukyuan**). A Japanese–Ainu relationship has been hypothesized, but evidence is scanty. There have been attempts to make more systematic comparisons between Japanese and Korean, but even here reliable sets of cognates are extremely small in number.

Being a mountainous country with numerous islands, Japan is an ideal setting for fostering language diversification. Indeed, Japanese is extremely rich in dialectal variations, and different dialects are often mutually unintelligible. However, Japan is linguistically completely unified by a uniform writing system and by the spread of the standardized speech, based on the Tokyo dialect, making oral communication possible among speakers of different dialects. The long literary history of Japanese, which dates back to the 8th century, finds its root in the borrowing of Chinese characters as a means of transcribing Japanese. Simplification of Chinese characters gave rise to two kinds of syllabary, or kana, i.e., *hiragana* and *katakana*. In addition to Chinese characters and the two types of syllabary, the Japanese writing system includes the Roman alphabet, which was introduced in the late 16th century by Portuguese and Spanish missionaries. All four types of writing systems have been retained today, and it is not unusual to see contemporary Japanese sentences written with a mixture of all of them (*see* **Japan: Writing System**).

The historical contacts with foreign cultures have left strong marks in the Japanese lexicon, which is characterized by a high percentage of loanwords. Roughly 60% of the Japanese vocabulary consists of loanwords of Chinese origin, a figure comparable to the proportion of Latinate words in the English vocabulary. Among the non-Chinese loanwords, or roughly 10% of Japanese vocabulary, English loans stand out, often replacing older loans from Portuguese, Spanish, and Dutch as well as even some Chinese loans.

Another salient feature of the Japanese lexicon is the presence of a large number of sound symbolic or mimetic words, which depict not only the sounds of natural objects and animals but also the manners of action and even states of mind. Thus, a dog barks *wan-wan* and it pours rain *zaa-zaa*; an old man walks *yobo-yobo* 'wobbly' and an old lady chatters *peya-kutya*; and your head aches *zuki-zuki* 'throbbingly,' your stomach hurts *tiku-tiku* 'stingingly,' and your nerves are irritated *ira-ira*.

Segmental phonology of Japanese – at least the speech of Tokyo and surrounding areas – is rather simple, with five vowel phonemes /a e i o u/ and 16 consonantal phonemes /p t k b d g s h z r m n w j N Q/. A noteworthy feature of Japanese segmental phonology is the distinction between a syllable and a mora. A mora is a unit that can be represented by one letter of kana (a Japanese pseudo-character used in syllabic writing). A word such as *simbun* 'newspaper' consists of two syllables, but a Japanese speaker further subdivides the word into four units, *si*, *m*, *bu*, and *n*, which correspond to the four letters of kana used in the written form. The consonantal archiphonemes /N/ and /Q/ correspond to moraic consonants seen in words such as *simbun* 'newspaper,' *simpai* 'worry,' *hakkiri* 'clearly,' and *kossori* 'stealthily,' wherein the first segments of the consonant clusters, *m*, *k*, and *s*, constitute moras. Since these moraic consonants are homorganic to the following consonants, there is no contrast among them other than in terms of the nasality feature. And since the only consonant that ends a word is the moraic nasal *n*, the words *simbun*, *simpai*, *hakkiri*, and *kossori* are phonemicized, respectively, as /siNbuN/, /siNpai/, /haQkiri/, and /koQsori/. The units of syllable and mora each play important roles in the accentual system.

Japanese is arguably a tone language, wherein the pitch pattern can be high followed by low (HL) or low followed by high (LH). Pitch height alone distinguishes minimal pairs such as *hasi* (HL) 'chopstick' versus *hasi* (LH) 'bridge' and *ame* (HL) 'rain' versus *ame* (LH) 'candy'; however, the Japanese accentual system is characteristically different from archetypical tone languages of the Chinese type, in which it is necessary to specify the tone for each syllable. In the case of (Tokyo) Japanese, once the place of pitch drop for a given word or a minor phonological phrase is specified, the pitch shape can be fully predicted. Thus, *hasi* 'chopstick' and *hasi* 'bridge' can be specified as /ha'si/ and /hasi'/, respectively, so that in the former, high pitch drops after *ha*, whereas in the latter, all moras are high except for the initial one. A pitch drop in the latter form is observed only when it is followed by another element, such as the nominative particle *ga*, as in the minor phrase /hasi'ga/ (LHL) 'bridge NOM.' The word *hasi* (LH) 'edge' is representable as /hasi/ without any accent; indeed, there is no pitch drop here, even in a phrase such as /hasi ga/ (LHH) 'edge NOM.' Pitch changes occur at mora boundaries, and the accent marker indicating the location of pitch drop is assigned to the unit of syllable. For example, the three-mora, two-syllable word *ganko* 'stubborn' contains the accent marker in the initial syllable, /ga'Nko/, surfacing with the pitch shape of HLL. There is no word in which the second mora of the

first syllable carries the accent, such as /gaN'ko/, which would presumably be pronounced as LHL or HHL if the syllable were the unit of tone inflection.

Japanese is an agglutinative language with primarily suffixing morphology. Both verbs and adjectives inflect for tense, but they are distinguished by different tense suffixes: *mi-ru* 'see-PRES' and *mi-ta* 'see-PAST' versus *utukusi-i* 'beautiful-PRES' and *utukusi-kat-ta* 'beautiful-EXPL-PAST' (EXPL = expletive). In addition to the inflecting adjectives, adjectival nominals are similar in meaning to adjectives but, like nominal predicates, call for tense-carrying copula (COP) in their predicative function: e.g., *kirei-da* 'pretty-COP-PRES' and *kirei-da-tta* 'pretty-COP-PAST.' These tense suffixes combine with various auxiliary-type suffixes, often resulting in a fairly long verbal complex: *ika-se-rare-ta-gara-na-i* (go-CAUS-PASS-DESI-show-NEG-PRES) 'do not show signs of wanting to be made to go.'

Japanese syntax is consistently head-final. The basic word order is subject-object-verb: *Taroo ga hon o yonda* (Taro NOM book ACC read-PAST) 'Taro read a book.' Postpositional particles are used instead of prepositions, as in the example, wherein the nominative and accusative (ACC) particles *ga* and *o*, respectively, mark the subject and the object. Modifiers precede the heads that they modify: *takai hon* (expensive book) 'expensive book' [*Taroo ga katta*] *hon* ([Taro NOM bought] book) 'the book Taro bought,' *Taroo no ie* (Taro GEN house) 'Taro's house,' *sono hon* (that book) 'that book,' *san-satu no hon* (three-CLASS GEN book) 'three books,' *hayaku hasiru* (quickly run) 'run quickly,' *tabe-tai* (eat-want) 'want to eat,' *Taroo yori kasikoi* (Taroo than smart) 'smarter than Taro.' Subordinating conjunctions occur after subordinate clauses, which in turn come before main clauses: [*Taroo ga kita*]-*node minna ga kaetta* ([Taro NOM came]-because everyone NOM went home) 'Because Taro came, everyone went home.'

One of the most important aspects of Japanese grammar has to do with the topic construction. The topic particle *wa* attaches to various nominals and adverbials, yielding topic sentences that contrast with nontopic sentences in the following manner:

Nontopic sentence (1):

(1) Taroo ga Ziroo no hon o yonde-iru
Taro NOM Jiro GEN book ACC read-be
'Taro is reading Jiro's book'

Topic sentences (2) and (3):

(2) Taroo wa Ziroo no hon o yonde-iru
Taro TOP Jiro GEN book ACC read-be
'Taro is such that he is reading Jiro's book'

(3) Ziroo no hon wa Taroo ga yonde-iru
Ziroo GEN book TOP Taro NOM read-be
'Jiro's book is such that Taro is reading it'

The basic difference between topic sentences and nontopic sentences is that the former are statements about certain things, represented by topic nominals, and the latter are statements describing the occurrences of events. For example, nontopic sentence (1) describes the event of Taro reading Jiro's book. On the other hand, topic sentences (2) and (3), respectively, describe something about 'Taro' and 'Jiro's book.' Thus, sentence (1) answers a question such as 'What is happening?,' whereas sentences (2) and (3), respectively, would be used in answering questions such as 'What is Taro doing?' and 'Where is Jiro's book?' In Japanese, it is the *wa*-marked topic nominal that accurately represents the traditional Western definition of subject as something that is being talked about. The *ga*-marked subject nominal, on the other hand, is more consonant with the other definition of subject, namely, that it expresses an actor or agent. In other words, in Japanese, the two notional definitions of subject are distributed over two distinct syntactic relations, whereas in English and other European languages, these largely converge on the single subject nominal.

Japanese has no agreement marker, but it freely omits pronouns. Thus, the following type of exchange is not uncommon:

(4) *Oo kita ka* (Oh came Q) 'Oh (you) came?'

(5) *Un kita* 'Yeah (I) came'

Of course, the omissions of pronouns are permitted only when they are recoverable from the context; one type of clue for the recovery is found in the honorific endings. Humbling forms together with the polite (POL) ending such as *Mair-imasu* (go.HUMBLE-POL) 'go' and *O-tazune simasu* (HON-visit do-POL) 'visit' indicate (a) that the subject is either a speaker or someone close to the speaker (by use of the humbling forms) and (b) that the addressee is someone worthy of respect (by use of the polite ending). On the other hand, honorific forms with the plain ending, such as *Oide-ni naru* (go.HON-ADV become.PLAIN) 'go' and *Otazune-ni naru* (HON-visit-ADV become.PLAIN) 'visit' indicate (a) that the subject is other than the speaker and is someone worthy of respect (honorific forms) and (b) that the addressee is someone close to the speaker, to whom the speaker is not obliged to show respect (plain ending). Thus, both the addressee axis (for polite endings) and the referent axis (for honorific and humbling endings) control honorific phenomena independently, though they often converge, as when the addressee is also the referent of the subject nominal (*see also* **Honorifics**).

Another clue that helps identify the nature of the speaker is sentence final particles, some of which are

different for male and female speakers. The final particle (PART) *wa* is a typical female form, whereas *zo* occurs in rough male speech. Since these discourse particles occur in intimate speech, an expression such as *Mair-imasu wa* (go.HUMBLE-POL PART) '(I will) go' indicates that the subject is the speaker (humble form), that the addressee is someone worthy of respect (polite ending), and that the speaker is a woman who is on intimate terms with the addressee (final particle).

Thus, Japanese, though it lacks agreement markers, has a number of grammatical features that not only indicate the nature of the subject but also index the social relationships between the speaker and the addressee and between the speaker and the nominal referent, as well as the gender of the speaker. These features, on the other hand, require the speaker of Japanese to predetermine the social relationships between the speaker and the addressee and the nominal referent, so that appropriate combinations of honorifics and discourse particles can be chosen. Japanese, in other words, is a highly context-sensitive language in which individual expressions encode various factors that make up conversational contexts in which they are embedded.

See also: Ainu; Altaic Languages; Austronesian Languages: Overview; Honorifics; Japan: Language Situation; Japan: Writing System; Japanese Lexicography; Korean; Ryukyuan.

Bibliography

Kuno S (1973). *The structure of the Japanese language.* Cambridge, MA: MIT Press.
Martin S E (1975). *A reference grammar of Japanese.* New Haven, CT: Yale University Press.
McCawley J D (1968). *The phonological component of the grammar of Japanese.* The Hague: Mouton.
Miller R A (1971). *Japanese and the other Altaic languages.* Chicago, IL: University of Chicago Press.
Shibatani M (1990). *The languages of Japan.* Cambridge: Cambridge University Press.

Japanese Lexicography

Y Tono, Meikai University, Chiba, Japan

There are three broad categories of dictionaries for the Japanese language: (a) Chinese–Japanese character dictionaries (*Kanwa Jiten*), (b) Japanese monolingual dictionaries (*Kokugo Jiten*), and (c) specialized dictionaries (*Tokushu Jiten*), which include thesauruses and terminology dictionaries. In terms of lexicographical development, it would be fair to divide the chronology into four main periods or stages: (1) the dawn of premodern lexicography (900 A.D.–1868 A.D.), (2) Westernization and rapid development of modern lexicography before World War II (1869–1945), (3) expansion of the dictionary market and projects (1946–1999), and (4) the electronic age (2000–).

The Dawn of Premodern Lexicography (900 A.D.–1868 A.D.)

Development of Kanwa Jiten

The oldest of the three categories of dictionaries is the Chinese–Japanese character dictionary. In the 7th century, Buddhism was introduced into Japan by way of Korea. Because there was no indigenous writing system in Japan at that time, the Buddhist monks had to learn how to read and write Chinese by consulting imported Chinese monolingual dictionaries, such as *Erya* (爾雅), *Shuowen Jiezi* (説文解字), and *Yupian* (玉篇). Gradually they compiled a list of Chinese characters with Japanese translations, which grew into a genre of *Kanwa Jiten* (漢和辞典; Chinese–Japanese character dictionaries). The first Chinese character dictionary produced in Japan was the *Shinsen Jikyou* (新撰字鏡), compiled by a Buddhist monk, Shouju, probably between 898 and 901 A.D. Following the *Shinsen Jikyou*, numerous Chinese–Japanese character dictionaries appeared. Among those produced from the Heian period (794 A.D.–1185) through the Edo period (1600–1868), four stand out: *Tenrei Banshou Meigi* (篆隷万象名義, 30 vols., early 9th century[1]) by a Buddhist monk, Kukai; *Ruiju Myogi Shou* (類聚名義抄, 10 vols, late 11th century, author unknown); and *Jikyou* (字鏡) and *Wagokuhen* (和玉篇) (dates and compilers unknown). The primary aim of these dictionaries was to provide information on how to read Chinese characters and their translations by using the Japanese writing system (*Manyogana*, a set of Chinese characters developed in the 8th century to

[1]An electronic version of *Tenrei Bansho Meigi* has been developed in Japan. See further details at http://homepage3.nifty.com/shikeda/kkg178.html.

different for male and female speakers. The final particle (PART) *wa* is a typical female form, whereas *zo* occurs in rough male speech. Since these discourse particles occur in intimate speech, an expression such as *Mair-imasu wa* (go.HUMBLE-POL PART) '(I will) go' indicates that the subject is the speaker (humble form), that the addressee is someone worthy of respect (polite ending), and that the speaker is a woman who is on intimate terms with the addressee (final particle).

Thus, Japanese, though it lacks agreement markers, has a number of grammatical features that not only indicate the nature of the subject but also index the social relationships between the speaker and the addressee and between the speaker and the nominal referent, as well as the gender of the speaker. These features, on the other hand, require the speaker of Japanese to predetermine the social relationships between the speaker and the addressee and the nominal referent, so that appropriate combinations of honorifics and discourse particles can be chosen. Japanese, in other words, is a highly context-sensitive language in which individual expressions encode various factors that make up conversational contexts in which they are embedded.

See also: Ainu; Altaic Languages; Austronesian Languages: Overview; Honorifics; Japan: Language Situation; Japan: Writing System; Japanese Lexicography; Korean; Ryukyuan.

Bibliography

Kuno S (1973). *The structure of the Japanese language.* Cambridge, MA: MIT Press.
Martin S E (1975). *A reference grammar of Japanese.* New Haven, CT: Yale University Press.
McCawley J D (1968). *The phonological component of the grammar of Japanese.* The Hague: Mouton.
Miller R A (1971). *Japanese and the other Altaic languages.* Chicago, IL: University of Chicago Press.
Shibatani M (1990). *The languages of Japan.* Cambridge: Cambridge University Press.

Japanese Lexicography

Y Tono, Meikai University, Chiba, Japan

There are three broad categories of dictionaries for the Japanese language: (a) Chinese–Japanese character dictionaries (*Kanwa Jiten*), (b) Japanese monolingual dictionaries (*Kokugo Jiten*), and (c) specialized dictionaries (*Tokushu Jiten*), which include thesauruses and terminology dictionaries. In terms of lexicographical development, it would be fair to divide the chronology into four main periods or stages: (1) the dawn of premodern lexicography (900 A.D.–1868 A.D.), (2) Westernization and rapid development of modern lexicography before World War II (1869–1945), (3) expansion of the dictionary market and projects (1946–1999), and (4) the electronic age (2000–).

The Dawn of Premodern Lexicography (900 A.D.–1868 A.D.)

Development of Kanwa Jiten

The oldest of the three categories of dictionaries is the Chinese–Japanese character dictionary. In the 7th century, Buddhism was introduced into Japan by way of Korea. Because there was no indigenous writing system in Japan at that time, the Buddhist monks had to learn how to read and write Chinese by consulting imported Chinese monolingual dictionaries, such as *Erya* (爾雅), *Shuowen Jiezi* (説文解字), and *Yupian* (玉篇). Gradually they compiled a list of Chinese characters with Japanese translations, which grew into a genre of *Kanwa Jiten* (漢和辞典; Chinese–Japanese character dictionaries). The first Chinese character dictionary produced in Japan was the *Shinsen Jikyou* (新撰字鏡), compiled by a Buddhist monk, Shouju, probably between 898 and 901 A.D. Following the *Shinsen Jikyou*, numerous Chinese–Japanese character dictionaries appeared. Among those produced from the Heian period (794 A.D.–1185) through the Edo period (1600–1868), four stand out: *Tenrei Banshou Meigi* (篆隷万象名義, 30 vols., early 9th century[1]) by a Buddhist monk, Kukai; *Ruiju Myogi Shou* (類聚名義抄, 10 vols, late 11th century, author unknown); and *Jikyou* (字鏡) and *Wagokuhen* (和玉篇) (dates and compilers unknown). The primary aim of these dictionaries was to provide information on how to read Chinese characters and their translations by using the Japanese writing system (*Manyogana*, a set of Chinese characters developed in the 8th century to

[1]An electronic version of *Tenrei Bansho Meigi* has been developed in Japan. See further details at http://homepage3.nifty.com/shikeda/kkg178.html.

be used as phonetic symbols to represent Japanese syllables, and *katakana* and *hiragana*, a system of 48 syllabic writing units for writing non-Chinese loan words and indigenous Japanese words, respectively).

Development of Kokugo Jiten

Because Chinese–Japanese character dictionaries were not able to describe Japanese indigenous words, there was a pressing need for Japanese monolingual dictionaries. One of the earliest extant examples of such a dictionary is the *Iroha jirui shou* (色葉字類抄), compiled by Tachibana no Tadakane between 1174 and 1181. It is a prototype of similar dictionaries in use today. Unlike Chinese character dictionaries, its entries are arranged according to sound (using *iroha* poetry) rather than Chinese characters. Nowadays most monolingual dictionaries are generally arranged on the basis of the standard *gojyuon* (50-sound) system. Notable examples of premodern monolingual dictionaries produced from the late Muromachi period (1333–1568) to the late Edo period include: *Setsuyoushu* (節用集; the first one was compiled in the mid-Muromachi period, with various subsequent versions appearing throughout the Edo period); *Wakun no shiori* (和訓栞, compiled by Tanikawa Kotosuga in the mid-Edo era; the first monolingual dictionary arranged in the *gojyuon* system); and *Gagen shoran* (雅言集覧, compiled by Ishikawa Masamochi from 1826–49; revised by Nakajima Hirotari in 1887; the first monolingual dictionary illustrating the meanings of old Japanese words with examples from authentic texts) (Kimura, 2002).

Early Western Influence

In the Azuchi-Momoyama (1573–1597) and the Edo (1600–1868) periods, lexicographical methods used in Europe were introduced through the Europeans who had contact with the Japanese under special circumstances. Among them were the Jesuit missionaries who compiled bilingual dictionaries such as *Rahonichi Jisho* (*Dictionaricum Latino-Lusitanicum ac Iaponicum,* 1595) and *Nippo Jisho* (*Vocabulario lingua de Iapam com adeclaraa em Portugues,* 1603). During the Edo period, the Netherlands was the only country that was allowed to have trade relations with Japan. People learned *Rangaku* (i.e., Dutch studies) in order to absorb Western culture. Inamura Sanpaku and others translated François Halma's Dutch–French dictionary, *Woordenboeck der Nederduitsche en Fransche Taalen* (1708), and compiled a Dutch–Japanese bilingual version called *Halma Wage* (波留麻和解, published in 1796; 80 000 entries). This version was called 'Edo Halma,' while the one translated later in 1816 by Hendrick Doeff, head of the Dutch trading post at Nagasaki, was called 'Nagasaki Halma.' The latter version was never printed, and students at the Tekijuku academy of Dutch studies in Osaka made their own manuscript copies. It was later revised by Katsuragawa Hoshu and published under the title *Oranda jii*.

It was not until the introduction of Western lexicography in the Meiji era that Japanese lexicography began its modern phase of development. One of the most influential dictionaries was James Curtis Hepburn's Japanese–English dictionary *Waei Gorin Shusei* (和英語林集成; 1867; Japanese–English: 20 772 entries; English–Japanese and: 10 030 entries). Due to its systematic description of Japanese words, this dictionary not only influenced other bilingual dictionaries (Japanese–German and Japanese–French, among others) but also had a great influence on the development of Japanese monolingual dictionaries. The third edition (Japanese–English: 35 618 entries; English–Japanese: 15 697 entries), which appeared in 1886, was especially famous as it introduced Hepburn's Romanization, a transliteration scheme for writing Japanese, for the very first time. This Romanization system became the standard in Japan and is still used today.

Westernization and Rapid Development of Modern Lexicography (1869–1945)

Meiji Restoration and Educational Reform

Two important watersheds in the development of Japanese dictionaries occurred in the late 19th and early 20th centuries. One was the introduction of foreign dictionaries and lexicographical methods in the late Edo period. The other was the system of compulsory education instituted in the early Meiji period (1868–1912). Because of this education reform, mass production of all types of dictionaries took off in Japan.

Among the dictionaries compiled during this period, *Genkai* (言海, edited by Otsuki Fumihiko) was the most influential and ambitious (Yamauchi, 1965). It originally started as a government-led lexicographical project. After the Meiji Restoration, the new government wished to have a new unabridged monolingual dictionary. This was influenced by linguistic nationalism in major Western countries, which had led to large-scale monolingual dictionary projects (e.g., the *Oxford English dictionary* in the United Kingdom, Grimm's *Deutsches Wörterbuch* in Germany, Webster's *An American dictionary of the English language* in the United States, and Littré's *Dictionnaire de la langue française* in France). The

government appointed around 10 Japanese language scholars for this work, and the first section of what was then called *Goi* (語彙) was published in 1871. The project had to be terminated, however, due to the lack of government funding and conflicting editorial policies among the team members. Otsuki Fumihiko was asked by the Education Ministry to take over this job in 1875 and spent 17 years completing the rest of the dictionary almost single-handedly. The first edition was published in 1889, and robust sales resulted in more than 500 printings by the beginning of the Showa period (around 1930).

Genkai was roughly based on Webster's dictionary in terms of its editorial policy. The criteria for selecting and organizing entries, especially for inflected forms, compound words and idioms, were much more refined and scientific than in previous dictionaries. It contained 39 103 entries, each of which consisted of information such as pronunciation, part of speech, word origin, senses, and citations. In particular, the structure of the entries became more scientific; definitions were ordered on historical principles and supplied with illustrative examples wherever possible. If a firm definition or explanation could not be fully established, a phrase such as *to iu ha ikaga* (i.e., 'suggested') was added after the description. Japanese linguistics was still in its infancy at that time, and the objective attitude adopted in the dictionary toward the description of the language became an important milestone in its development. *Genkai* was highly regarded and welcomed by scholars and became the most authoritative dictionary during the Meiji era, with a high academic standing among Japanese scholars and the general population alike. Most modern Japanese dictionaries are still influenced by *Genkai* to some extent.

After *Genkai* set the standard for medium-sized Japanese monolingual dictionaries, many attempts were made to match or even surpass this work. Among these attempts were *Jirin* (辞林, 1907; revised as *Kojirin* in 1925) by Kanazawa Shozaburo, *Jien* (辞苑, 1935) by Shinmura Izuru, and *Jikai* (辞海, 1952) by Kindaichi Kyosuke.

The publication of *Genkai* encouraged publishers to provide more thorough descriptions of the Japanese language, which led to some attempts to publish more comprehensive, unabridged dictionaries. For instance, *Genkai* was later revised and updated by Otsuki's successor, Okubo Toshio, and other Japanese language scholars at the University of Tokyo. This enlarged version was entitled *Dai-Genkai* (大言海, 4 vols, 1932–37; c. 80 000 entries). Other unabridged dictionaries in the early 20th century include *Dai Nihon Kokugo Jiten* (大日本国語辞典, 4 vols, edited by Matsui Kanji & Ueda Kazutoshi, 1915–19; c. 220 000 entries; most examples with references), *Nihon Daijiten Gensen* (日本大辞典言泉, by Haga Yaichi, 1921–29; many encyclopedic entries), and *Daijiten* (大辞典, 24 vols, by Heibonsha, 1934–36; encyclopedic in nature).

Expansion of the Dictionary Market and Projects (1946–1999)

After World War II, Japan needed some time to recover from its chaotic state. In the 1950s, however, rapid recovery and economic growth encouraged dictionary publishers to produce dictionaries capitalizing on their prewar heritage. *Jien* (1935), for example, was revised and updated and renamed *Kojien* (広辞苑, edited by Shinmura Izuru, 1955; 200 000 entries), which became a best-selling medium-sized monolingual dictionary for a long time.

Major dictionaries published in the latter half of the Showa period are:

- *Jikai* (辞海, edited by Kindaichi Kyosuke, 1952; 130 000 entries)
- *Shincho Kokugo Jiten* (新潮国語辞典, edited by Yamada Toshio, 1965; 130 000 entries)
- *Nihon Kokugo Daijiten* (日本国語大辞典, edited by Kankokai, 1972–1976; 450 000 entries)
- *Gakken Kokugo Daijiten* (学研国語大辞典, edited by Kindaichi Haruhiko and Ikeda Yosaburo, 1978; 100 000 entries)
- *Kadokawa Kokugo Daijiten* (角川国語大辞典, edited by Tokieda Motoki *et al.*, 1982; 155 000 entries)
- *Gensen* (言泉, edited by Hayashi Oki, 1986; 150 000 entries)
- *Daijirin* (大辞林, edited by Matumura Akira, 1988; 220 000 entries)
- *Kodansha Color Nihongo Daijiten* (講談社カラー版日本語大辞典, edited by Umesao Tadao *et al.*, 1989; 175 000 entries).

One of the largest and most comprehensive dictionaries is the 20-volume *Nihon Kokugo Daijiten* completed by Shogakukan Publishing Co. Ltd. in 1976. This project was ambitious in the sense that they designed the dictionary with the *Oxford English dictionary* as their model, with not only definitions for the entries but also information on their origins and citations showing a chronological range of meanings, from the time the word entered the language to the present. It contains 450 000 entries, with 750 000 illustrative examples from more than 30 000 references. Detailed descriptions are provided for pronunciation variants (regional differences in accents), and more than one etymology is given for words of uncertain origin. Definitions in this dictionary are ordered according to word-sense derivation, and citations are

given with clear references. It is also noteworthy that this dictionary makes reference to all the major Japanese dictionaries published before the Meiji period to indicate whether the given word was already listed in those dictionaries (Kurashima, 1997).

Nihon Kokugo Daijiten was welcomed enthusiastically by the Japanese public as an authoritative monolingual dictionary. The second edition was published in 2001 (500 000 entries; 1 million citations) and is the biggest monolingual Japanese dictionary published so far.

During this period, hundreds of monolingual Japanese dictionaries were produced, catering to all the different sectors and niches of market. The major trends in modern Japanese lexicography can be summarized as follows.

1. Attention has shifted away from mere explanations of difficult terms to more careful linguistic descriptions of ordinary words. Unabridged monolingual dictionaries in Japan used to cover classical words and phrases more fully and provided rather sparse treatment of modern usage and vocabulary. Even the most popular dictionary, *Kojien* (1995), had ordered word senses on historical principles. Its rival, *Kojirin*, on the other hand, lists the most recent senses first, beginning with modern usages and working backward to the word's origins. Nowadays, even small, compact dictionaries deal with everyday words in some detail.
2. Illustrative examples are now also taken from modern texts. Citations used to be from classical literature and were given only for the more difficult terms. Contemporary dictionaries use more examples from contemporary texts.
3. There has been a growing awareness that the needs of the end user should be taken into account in designing a dictionary. This is shown in the fact that monolingual dictionaries designed especially for schoolchildren now focus on more productive knowledge of vocabulary (everyday expressions, collocations, word choice, etc.).

This period was also characterized by the publication of more Chinese-character dictionaries and various types of terminological dictionaries. Of the numerous *kanwa jiten* published for general use, the following achieved a particularly wide circulation: *Kanwa Daijiten* (1903), jointly compiled by Shigeno Yasutsugu, Mishima Takashi, and Hattori Unokichi (1867–1939), and *Daijiten* (1917), by Ueda Kazutoshi. The most comprehensive of the modern kanwa jiten is the *Dai Kanwa Jiten*, compiled by Morohashi Tetsuji and published in 1955–60.

By the year 2000, there were already more than 2500 terminological dictionaries available. Some are general dictionaries with a special emphasis on certain technical fields, whereas others are more like lists of terminology along with translation equivalents and brief comments. There are dictionaries that deal exclusively with recently coined words, slang, trade terms, place and personal names, recently introduced loan words, acronyms, sports and leisure activities, and so on. Field-specific dictionaries exists for such subjects as mathematics, physics, chemistry, computer science, linguistics, Japanese literature, world literature, music, history, geography, medicine, social work, transportation, architecture, communication, law, economics, management, business, education, the performing arts, and movies.

The Electronic Age (from 1985)

In the field of natural language processing, a nine-year project called the EDR (Electronic Dictionary Research) (1986–1994) was launched. It aimed to produce dictionaries for human language technology. The EDR electronic dictionary is a machine-tractable dictionary that catalogues the lexical knowledge of Japanese and English (a word dictionary, a bilingual dictionary, and a co-occurrence dictionary), and has unified thesaurus-like concept classifications (a concept dictionary) with corpus databases (the EDR corpus). It is now widely available for research and commercial purposes.

As computer technology advanced in the late 20th century, Japanese lexicography gradually moved into a new age of electronic publishing. The first CD-ROM dictionary was *Saishin kagaku yogo jiten* (最新科学用語辞典, *A new dictionary of scientific terms*, Sanshusha) in 1985. In 1987, Iwanami Press published the CD-ROM version of *Kojien*, the most popular monolingual dictionary in Japan, which triggered a rapid growth of the electronic dictionary market. In 1991, Fujitsu, Sony, Dai Nippon Printing Co. Ltd, and Iwanami Press, all involved in the production of *Kojien*, founded the EPWING Consortium, and most publishers involved in electronic dictionary publishing joined it. Until the Internet and Web technology began to flourish, this EPWING Consortium facilitated the appearance of various types of electronic dictionaries on the market. Sony's *Denshi book* (Electronic Book DD–25) became very popular because it contained several dictionaries from different publishers on the same diskette. More recently, websites such as Yahoo and Infoseek have begun to provide free Web dictionary services, in which major unabridged monolingual dictionaries such as *Dai jirin* (大辞林 233 000 entries) and *Dai jisen* (大辞泉 220 000 entries) can be accessed via the Web. Also, pocket electronic dictionaries

produced by Casio, Seiko, Sharp, and Sony all contain several Japanese monolingual as well as Chinese character dictionaries. The electronic dictionary market is rapidly growing in Japan, and the market share will soon exceed that of paper dictionaries. The most popular type of these electronic dictionaries is the all-in-one type, which contains more than 30 different titles, including all types of Japanese dictionaries (monolingual, Chinese character, thesaurus, old Japanese, idioms, proverbs, and loan words), different English dictionaries (monolingual, Japanese–English, English–Japanese), encyclopedias, and some specialized dictionaries such as dictionaries of medicine and new words.

See also: Japan: Writing System; Japanese.

Bibliography

Hausmann F J *et al.* (eds.) (1989–1991) (3 vols). *Dictionaries.*

Inukai M (1999). *Kindai Kokugo Jisho Hensanshi no Kisoteki Kenkyu.* (*Foundations of modern Japanese lexicography: historical perspectives*). Tokyo: Kazama Shobo.

Japan: an illustrated encyclopedia. Kodansha.

Jiten K (ed.) (1996). *Nihon no Jisho no Ayumi* (*A history of Japanese dictionaries*). Jiten Kyokai.

Kimura A (2002). *Chusei Jisho no Kisoteki Kenkyu.* (*A basic study of Middle-Japanese dictionaries*). Tokyo: Kyuko Shoin.

Kurashima N (1997). *"Kokugo" to "Kokugo Jiten" no Jidai* (2 vols). Tokyo: Shogakukan.

Matsui T (1990). *Kindai Kango Jisho no Seiritsu to Tenkai.* (*Development of modern Chinese-character dictionaries*). Tokyo: Kasama Shoin.

Yamada T (1981). *Kindai Kokugo Jiten no Ayumi.* (*A history of modern Japanese dictionaries*). Tokyo: Sanseido.

Yamada T (2003). *Nihon no Kotoba to Kojisho* (*Japanese words and old dictionaries*). Tokyo: Sanseido.

Yamauchi S (1967). *Shosetsu Genkai* (*'Genkai' a novel*). Tokyo: Shinbisha.

Jargon

K Allan, Monash University, Victoria, Australia

The word *jargon* probably derives from the same source as *gargle*, namely Indo-European **garg*- meaning throat, and it originally referred to any noise made in the throat. In Middle English it was generally used to describe the chattering of birds, or human speech that sounded as meaningless as the chattering of birds. It was used (contemptuously) to refer to trade languages such as *Chinook Jargon* (q.v.). Jane Austen used *jargon* in the sense 'cliché' (*Sense and sensibility*, Ch.18). Today,

> Jargon is the language peculiar to a trade, profession, or other group; the language used in a body of spoken or written texts dealing with a circumscribed domain in which speakers share a common specialized vocabulary, habits of word usage, and forms of expression.

This definition includes what some scholars call 'specialist' or 'technical' language, 'restricted' language (Firth, 1968: 98), 'sublanguage' (Kittredge and Lehrberger, 1982), and others 'register' (e.g., Zwicky and Zwicky, 1982; Wardhaugh, 1986).

Jargons differ from one another grammatically and sometimes phonologically or typographically, as can be seen by comparing a statement of some of the requirements on the cricket field with the two-line excerpt from a knitting pattern, then the wedding invitation that follows it, and all of these with the excerpt from a Wordsworth poem and a text message version of that (such as might be conveyed using the SMS facility on a mobile phone).

> A fast-medium right arm inswing bowler needs two or three slips, a deep third man, a gully, a deepish mid-off, a man at deep fine leg and another at wide mid-on.

> Cast on 63 sts: Knit 6 rows plain knitting.
> 7th row: K4, wl. fwd. K2 tog to the last 3 sts. K3.


Sue and Graeme Cannon
together with
Anne and Bob Ervin
extend a warm invitation to
Natalie and friend
to share with us the joyous occasion of the wedding of
Vikki and Ken
in the gardens of Nathania Springs, Monbulk
at 4.45pm on Sunday 10th May 1992
Following the ceremony, a celebration will be held
in their honour at the adjoining reception centre.

R.S.V.P. 1st May, 1992

produced by Casio, Seiko, Sharp, and Sony all contain several Japanese monolingual as well as Chinese character dictionaries. The electronic dictionary market is rapidly growing in Japan, and the market share will soon exceed that of paper dictionaries. The most popular type of these electronic dictionaries is the all-in-one type, which contains more than 30 different titles, including all types of Japanese dictionaries (monolingual, Chinese character, thesaurus, old Japanese, idioms, proverbs, and loan words), different English dictionaries (monolingual, Japanese–English, English–Japanese), encyclopedias, and some specialized dictionaries such as dictionaries of medicine and new words.

See also: Japan: Writing System; Japanese.

Bibliography

Hausmann F J *et al*. (eds.) (1989–1991) (3 vols). *Dictionaries.*

Inukai M (1999). *Kindai Kokugo Jisho Hensanshi no Kisoteki Kenkyu.* (*Foundations of modern Japanese lexicography: historical perspectives*). Tokyo: Kazama Shobo.

Japan: an illustrated encyclopedia. Kodansha.

Jiten K (ed.) (1996). *Nihon no Jisho no Ayumi* (*A history of Japanese dictionaries*). Jiten Kyokai.

Kimura A (2002). *Chusei Jisho no Kisoteki Kenkyu.* (*A basic study of Middle-Japanese dictionaries*). Tokyo: Kyuko Shoin.

Kurashima N (1997). *"Kokugo" to "Kokugo Jiten" no Jidai* (2 vols). Tokyo: Shogakukan.

Matsui T (1990). *Kindai Kango Jisho no Seiritsu to Tenkai.* (*Development of modern Chinese-character dictionaries*). Tokyo: Kasama Shoin.

Yamada T (1981). *Kindai Kokugo Jiten no Ayumi.* (*A history of modern Japanese dictionaries*). Tokyo: Sanseido.

Yamada T (2003). *Nihon no Kotoba to Kojisho* (*Japanese words and old dictionaries*). Tokyo: Sanseido.

Yamauchi S (1967). *Shosetsu Genkai* (*'Genkai' a novel*). Tokyo: Shinbisha.

Jargon

K Allan, Monash University, Victoria, Australia

The word *jargon* probably derives from the same source as *gargle*, namely Indo-European **garg-* meaning throat, and it originally referred to any noise made in the throat. In Middle English it was generally used to describe the chattering of birds, or human speech that sounded as meaningless as the chattering of birds. It was used (contemptuously) to refer to trade languages such as *Chinook Jargon* (q.v.). Jane Austen used *jargon* in the sense 'cliché' (*Sense and sensibility*, Ch.18). Today,

> Jargon is the language peculiar to a trade, profession, or other group; the language used in a body of spoken or written texts dealing with a circumscribed domain in which speakers share a common specialized vocabulary, habits of word usage, and forms of expression.

This definition includes what some scholars call 'specialist' or 'technical' language, 'restricted' language (Firth, 1968: 98), 'sublanguage' (Kittredge and Lehrberger, 1982), and others 'register' (e.g., Zwicky and Zwicky, 1982; Wardhaugh, 1986).

Jargons differ from one another grammatically and sometimes phonologically or typographically, as can be seen by comparing a statement of some of the requirements on the cricket field with the two-line excerpt from a knitting pattern, then the wedding invitation that follows it, and all of these with the excerpt from a Wordsworth poem and a text message version of that (such as might be conveyed using the SMS facility on a mobile phone).

> A fast-medium right arm inswing bowler needs two or three slips, a deep third man, a gully, a deepish mid-off, a man at deep fine leg and another at wide mid-on.

> Cast on 63 sts: Knit 6 rows plain knitting.
> 7th row: K4, wl. fwd. K2 tog to the last 3 sts. K3.

Sue and Graeme Cannon
together with
Anne and Bob Ervin
extend a warm invitation to
Natalie and friend
to share with us the joyous occasion of the wedding of
Vikki and Ken
in the gardens of Nathania Springs, Monbulk
at 4.45pm on Sunday 10th May 1992
Following the ceremony, a celebration will be held
in their honour at the adjoining reception centre.

R.S.V.P. 1st May, 1992

Earth has not anything to shew more fair:	erth nt a thng so brill
Dull would he be of soul who could pass by	hes dul v soul pssng by
A sight so touching in its majesty.	sght of mjstic tch.
(Wordsworth *Upon Westminster Bridge*)	(Peter Finch *N Wst Brdg*, a text message version of the Wordsworth lines from http:// .guardian.co./uk//// |0,12241,785819,00.html in September2002)

A jargon is identified by one or more of the following criteria.

i. **Lexical markers.**
   a. Vocabulary specialized for use in some particular domain (the subject matter of a jargon). The lexical relations among specialized vocabulary will reflect the accepted taxonomies within the domain (e.g., forms, varieties, species, genera, families, orders, classes in biology).
   b. Idioms and abbreviations (e.g., in telecommunications, *DNA* 'does not answer', *MBC* 'major business customer', *HC&F* 'heat coil and fuse', *LIBFA* 'line bearer fault analysis'; in linguistics, *Noun Phrase* and NP; in logic, *if and only if* and *iff*; in biology, ♂ and ♀).

ii. **Syntactic markers** such as
   a. imperatives in recipes and knitting patterns;
   b. large numbers of impersonal passives in reports of scientific experiments (e.g., *It was observed that . . .*);
   c. full noun phrases in place of pronouns in legal documents (e.g., *A term of a sale shall not be taken to exclude, restrict, or modify the application of this Part unless the term* [not 'it'] *does so expressly or is inconsistent with that provision*).

iii. **Presentational markers.**
   a. Prosodic (voice quality, amplitude, rhythm, etc.) and paralinguistic and/or kinesic (gaze, gesture, etc.) characteristics within a spoken medium, typographical conventions within a written medium; e.g., a hushed tone and minimal kinesic display is more frequently expected in funeralese than in football commentary or anecdote; in mathematics, {a,b}, <a,b>, and (a,b) will normally have different and conventionally prescribed interpretations; in linguistics, language expressions that are mentioned rather than used are usually italicized.
   b. Format in which a text is presented; this is particularly evident in the written medium, as can be seen from the preceding four examples.

Jargon has two functions:

1. to serve as a technical or specialist language;
2. to promote in-group solidarity, and to exclude as out-groupers those people who do not use the jargon.

To the initiated, jargon is efficient, economical, and even crucial in that it can capture distinctions not made in the ordinary language. Linguists, for instance, redefine everyday terms such as *sentence, word, syllable,* and *grammar* and add a number of new terms to overcome imprecision and to distinguish things that nonlinguists ignore and, in consequence, ordinary language lacks terms for. They distinguish between *grammatical, orthographic*, and *phonological words* as well as introducing terms such as *lex, lexeme, morph,* and *morpheme* to capture additional distinctions.

Jargons, like bureaucratese (the language of government and corporate offices), have two motivations. One, shared with criminal jargon and slang, is the exclusion of out-groupers. The second motivation is to augment the bureaucrat's self-image by using a Graeco-Latinate lexicon, which (to the outsider, at least) obfuscates the commonplace (and often trivial) and endows it with gravity; this achieves a double-whammy by also mystifying and intimidating the clientele. We rarely uncover such blatant lexical substitution as

> [*Emendation to the traffic plan for a London borough*] Line 5. Delete 'Bottlenecks', insert 'Localised Capacity Deficiencies'. (Quoted in Cutts and Maher, 1984: 45)

No wonder Charles Dickens referred to Whitehall as the Circumlocution Office.

Because it is founded on a common interest, the most remarkable characteristic of a jargon is its specialized vocabulary and idiom. Although jargons facilitate communication among in-groupers on the one hand, on the other they erect communication barriers to keep out-groupers out. It is, of course, out-groupers who find jargon 'abounding in uncommon or unfamiliar words,' and therefore "unintelligible or meaningless talk or writing; gibberish" (*Macquarie Dictionary,* 1991). If the out-grouper is sufficiently rancorous she or he might also conclude the jargon is 'debased, outlandish or barbarous.' For some jargons such as legalese or linguisticalese, the in-group is fairly well defined. For others, such as stock market reports and games such as bridge, football, and cricket, there is barely an in-group at all except

among professionals. For language dealing with death (outside the funeralese of the bereavement industry), birth notices, and recipes, there is even less of an in-group – though there is a special vocabulary and there are conventional patterns of expression for all of these. The facts are clear: where in-groupers are associated with a particular trade or profession, they constitute a well-defined group; they are somewhat less well defined by a common recreational interest; and the most ill-defined in-groups are those defined merely by a temporary interest in the topic of the jargon – such as those members of the public who publish birth notices.

It follows that every text/discourse, including social chitchat, is jargon of one kind or another and that everybody uses at least one jargon. In fact, almost everybody controls several jargons and often many. It is often the case that an expert in one domain (an in-grouper with respect to its jargon) needs to explain something within the domain to a novice outside the domain (an out-grouper) with minimal use of jargon. An example would be where a lawyer needs to explain some point of law to a client, or a doctor needs to explain a medical condition to a patient, and educationists most all of the time. Any jargon is in constant contact with others, and it should be fairly obvious that jargons are not discrete from one another: all of them borrow from language that is common to other jargons. The following is an excerpt from a chatroom interchange replete with chatroom jargon such as *lol* (the acronym from "laughing out loud," which is used here more like a grin in *f2f* 'face to face' conversation as a mark of empathy), *j/k* 'just kidding', and emoticon smiles :), :-) and wicked smile >:); there is also jargon appropriate to in-group discussion of computer hardware, but note the divergence into less esoteric matters than disk drives and controllers. This mixing of jargons is very common in regular conversational discourse.

<Darkman-X> so wait becuz i have the onboard raid controller i could fit 2 hdd's on the raid controllers and 2 on the ide controllers?

<mark> RAID is just a standard of combining different drives.. you can make a raid out of scsi drives along side ide drives if you want

<mark> you could make a raid out of a hard drive and a ram drive

<mark> just forget about the raid stuff

<Darkman-X> lol

<mark> and picture it as another ide controller

<mark> thats all it is unles syou setup the raid stuff

<Darkman-X> so wait i can have 2 hdd's on the raid, 2 on teh primary ide and then 2 drives on the secondary ide?

<mark> i guess.. tried reading the manual?

<Darkman-X> lol manual?

<mark> :)

<mark> didnt your motherboard come with any papers

<Darkman-X> iz that that book that says A7V333 on it?

<[RaW]> yes

<Darkman-X> lol the one that i'm using to prop up my comp table?

<[RaW]> probably

<Darkman-X> whoops

<Darkman-X> :-)

<Darkman-X> j/k

<JoHn> lol

<[RaW]> yur supposed to use your school books for that dummy

<Darkman-X> ummm no i dun wanna look at em

<Darkman-X> they're evil

<[RaW]> keep your motherboard manual close to your heart

<[RaW]> >:)

<Darkman-X> lol

(*Sic*. Logged August 29, 2002)

A jargon cannot be precisely defined because the boundaries of any one particular jargon are impossible to draw nonarbitrarily.

It is impossible to taboo jargon. Jargon cannot be translated into 'ordinary English' (or whatever language) because there is no such thing. Changing the jargon alters the message: a speaker simply cannot exchange *faeces* for *shit* or *terrorist* for *freedom-fighter* or even *bottlenecks* for *localised capacity deficiencies* without changing the connotations of the message the person intends to convey. There is no convenient substitute for some jargon: to replace legalese *defendant* with *a person against whom civil proceedings are brought* is communicatively inefficient. It would be inappropriate for a lawyer not to use jargon when creating a legal document: that is exactly what legalese is for. Legal language is difficult because laws are complex and not because lawyers try to obfuscate. Similar remarks apply to other well-motivated uses of jargon.

For additional discussion, see Partridge, 1952; Hudson, 1978; Gowers, 1987; Green, 1987; Lutz, 1989; Asprey, 1991; Nash, 1993; Burke and Porter, 1995; Lutz, 1999; Allan, 2001; Bell, 2003; Allan and Burridge, 2006.

See also: Register: Overview.

Bibliography

Allan K (2001). *Natural language semantics.* Oxford & Malden, MA: Blackwell.

Allan K & Burridge K (2006). *Forbidden words: Taboo and the censoring of language.* Cambridge: Cambridge University Press.

Asprey M M (1991). *Plain language for lawyers.* Sydney: Federation Press.
Bell A (2003). *Your sharemarket jargon explained: tricks, traps and insider hints.* Milsons Point: Random House.
Burke P & Porter R (eds.) (1995). *Languages and jargons: contributions to the social history of language.* Cambridge: Polity Press.
Cutts M & Maher C (1984). *Gobbledygook.* London: George Allen & Unwin.
Firth J R (1968). 'Descriptive linguistics and the study of English.' In Palmer F R (ed.) *Selected papers of J. R. Firth, 1952–59.* Bloomington: Indiana University Press. 96–113. [First published 1956].
Gowers E (1987). *The complete plain words.* Revised by Sidney Greenbaum and Janet Whitcut. Harmondsworth: Penguin.
Green J (1987). *Dictionary of jargon.* London: Routledge & Kegan Paul.
Hudson K (1978). *The jargon of the professions.* London: Macmillan.
Kittredge R & Lehrberger J (eds.) (1982). *Sublanguage: studies of language in restricted semantic domains.* Berlin: De Gruyter.
Lutz W (1989). *Doublespeak: from "revenue enhancement" to "terminal living". How government, business, advertisers, and others use language to deceive you.* New York: Harper and Row.
Lutz W (1999). *Doublespeak defined: cut through the bull**** and get to the point.* New York: Harper Resource.
Macquarie dictionary (rev. 2nd edn.). McMahons Point: Macquarie Library.
Nash W (1993). *Jargon: its uses and abuses.* Oxford: Blackwell.
Partridge E (1952). *Introduction to chamber of horrors: A glossary of official jargon both English and American.* 'Vigilans' (ed.). London: Andre Deutsch.
Wardhaugh R (1986). *An introduction to sociolinguistics.* Oxford: Basil Blackwell.
Zwicky A & Zwicky A (1982). 'Register as a dimension of linguistic variation.' In Kittredge R & Lehrberger J (eds.) *Sublanguage: studies of language in restricted semantic domains.* Berlin: De Gruyter. 213–218.

Jarring, Gunnar (1907–2002)

B Sigurd, Lund, Sweden

Gunnar Valfrid Jarring, Turkologist and diplomat, was born on October 12, 1907 and died on May 28, 2002. He studied German and Scandinavian languages at Lund University, but it was during his comparative Indo-European studies with Herbert Pettersson and Slavic studies with Sigurd Agrell that he became interested in the Turkish languages. Turkish, in particular East Turkish, was represented by Gustaf Raquette, docent at Lund since 1924. Raquette and his missionary colleagues at Kashgar in East Turkestan (currently Xinjiang) were willing to receive Jarring, and he traveled by train and on horseback via Moscow over Pamir to Kashgar. There he started to collect linguistic data and manuscripts, which formed the basis of his later work. The collection was stored at the Lund University library. Unfortunately, Jarring contracted Typhus fever, which decreased his working capacity in Kashgar.

The missionary station in Kashgar (closed in 1938 for political reasons) had its own printing works, and Jarring published a catalogue of all the works published in *Prints from Kashgar* (1991). The missionaries suggested orthographic rules and norms that were important for the development of East Turkish. Jarring has described his journeys to and stay at Kashgar in the popular book *Åter till Kashgar* (1979; English version, 1986).

Jarring spent some months in Berlin during 1928 and participated in the Turkology seminar there headed by the famous Turkologist Willy Bang. The discussions convinced Jarring of the importance of raw concrete data and of living languages, which he always focused on in his works rather than on theoretical discussions. Jarring became a well-established Turkologist with colleagues and friends in many countries, which was important to him even in his diplomatic service.

After returning to Sweden, Jarring completed his thesis *Studien zu einer osttürkischen Lautlehre*, which he defended at Lund in 1933. The book is an impressive survey with evaluations of earlier works, including phonology and grammar. Jarring was aware of the shortcomings of earlier transcription systems and was a pioneer in using the new phonetic alphabet APA, a predecessor of the International Phonetic Alphabet (IPA). The thesis also includes some illustrative texts with translations and comments. He became docent in Turkology at Lund the same year (and honorary professor in 1986).

Jarring focused on field data and folkloristic texts, as illustrated by *Materials to the knowledge of Eastern Turki: tales, poetry, proverbs, riddles,*

Asprey M M (1991). *Plain language for lawyers.* Sydney: Federation Press.
Bell A (2003). *Your sharemarket jargon explained: tricks, traps and insider hints.* Milsons Point: Random House.
Burke P & Porter R (eds.) (1995). *Languages and jargons: contributions to the social history of language.* Cambridge: Polity Press.
Cutts M & Maher C (1984). *Gobbledygook.* London: George Allen & Unwin.
Firth J R (1968). 'Descriptive linguistics and the study of English.' In Palmer F R (ed.) *Selected papers of J. R. Firth, 1952–59.* Bloomington: Indiana University Press. 96–113. [First published 1956].
Gowers E (1987). *The complete plain words.* Revised by Sidney Greenbaum and Janet Whitcut. Harmondsworth: Penguin.
Green J (1987). *Dictionary of jargon.* London: Routledge & Kegan Paul.
Hudson K (1978). *The jargon of the professions.* London: Macmillan.
Kittredge R & Lehrberger J (eds.) (1982). *Sublanguage: studies of language in restricted semantic domains.* Berlin: De Gruyter.
Lutz W (1989). *Doublespeak: from "revenue enhancement" to "terminal living". How government, business, advertisers, and others use language to deceive you.* New York: Harper and Row.
Lutz W (1999). *Doublespeak defined: cut through the bull**** and get to the point.* New York: Harper Resource.
Macquarie dictionary (rev. 2nd edn.). McMahons Point: Macquarie Library.
Nash W (1993). *Jargon: its uses and abuses.* Oxford: Blackwell.
Partridge E (1952). *Introduction to chamber of horrors: A glossary of official jargon both English and American.* 'Vigilans' (ed.). London: Andre Deutsch.
Wardhaugh R (1986). *An introduction to sociolinguistics.* Oxford: Basil Blackwell.
Zwicky A & Zwicky A (1982). 'Register as a dimension of linguistic variation.' In Kittredge R & Lehrberger J (eds.) *Sublanguage: studies of language in restricted semantic domains.* Berlin: De Gruyter. 213–218.

Jarring, Gunnar (1907–2002)

B Sigurd, Lund, Sweden

Gunnar Valfrid Jarring, Turkologist and diplomat, was born on October 12, 1907 and died on May 28, 2002. He studied German and Scandinavian languages at Lund University, but it was during his comparative Indo-European studies with Herbert Pettersson and Slavic studies with Sigurd Agrell that he became interested in the Turkish languages. Turkish, in particular East Turkish, was represented by Gustaf Raquette, docent at Lund since 1924. Raquette and his missionary colleagues at Kashgar in East Turkestan (currently Xinjiang) were willing to receive Jarring, and he traveled by train and on horseback via Moscow over Pamir to Kashgar. There he started to collect linguistic data and manuscripts, which formed the basis of his later work. The collection was stored at the Lund University library. Unfortunately, Jarring contracted Typhus fever, which decreased his working capacity in Kashgar.

The missionary station in Kashgar (closed in 1938 for political reasons) had its own printing works, and Jarring published a catalogue of all the works published in *Prints from Kashgar* (1991). The missionaries suggested orthographic rules and norms that were important for the development of East Turkish. Jarring has described his journeys to and stay at Kashgar in the popular book *Åter till Kashgar* (1979; English version, 1986).

Jarring spent some months in Berlin during 1928 and participated in the Turkology seminar there headed by the famous Turkologist Willy Bang. The discussions convinced Jarring of the importance of raw concrete data and of living languages, which he always focused on in his works rather than on theoretical discussions. Jarring became a well-established Turkologist with colleagues and friends in many countries, which was important to him even in his diplomatic service.

After returning to Sweden, Jarring completed his thesis *Studien zu einer osttürkischen Lautlehre*, which he defended at Lund in 1933. The book is an impressive survey with evaluations of earlier works, including phonology and grammar. Jarring was aware of the shortcomings of earlier transcription systems and was a pioneer in using the new phonetic alphabet APA, a predecessor of the International Phonetic Alphabet (IPA). The thesis also includes some illustrative texts with translations and comments. He became docent in Turkology at Lund the same year (and honorary professor in 1986).

Jarring focused on field data and folkloristic texts, as illustrated by *Materials to the knowledge of Eastern Turki: tales, poetry, proverbs, riddles,*

ethnological and historical texts from the southern parts of Eastern Turkestan (1946–1991) and *The Moen collection of Eastern Turki (New Uighur) popular poetry* (1996). He treated the demographic situation in *On the distribution of Turk tribes in Afghanistan: an attempt at a preliminary classification* (1939). As an example of his religious observations one may mention *Culture clash in central Asia: Islamic views on Chinese theatre: Eastern Turki texts* (1991). During his stay as ambassador in Moscow from 1964 to1973, Jarring became interested in the history of Swedish–Russian relations, including the interesting destinies of the prisoners from the wars of Charles XII (the Caroleans). Jarring also examined the place-names noted by his famous compatriot Sven Hedin and published *Central Asian Turkic place-names: Lop Nor and Tarim area: an attempt at classification and explanation based on Sven Hedin's diaries* (1997).

Jarring became engaged in Swedish diplomatic service in 1941 and served in Teheran, New Delhi, Washington, and Moscow. He was the Swedish representative in the United Nations (UN) security council in 1957–1958. He was the UN secretary's special representative for the Middle East in 1967–1990 and served as mediator in the Kashmir conflict and in the conflict after the Six-Day War between Egypt and Israel. His name was also suggested for the post of secretary of the United Nations. He has reported about his life and experience in several books, e.g., *Memoarer 1939–1952* (1981).

As an international diplomat Jarring was known as 'The Clam,' and it was said about him that 'few could keep silent in as many languages as he.' In spite of his many important diplomatic tasks he managed to work continually in his special linguistic and ethnographic fields. His list of publications includes over 1000 items. He continued to extend his *Eastern Turki-English Dialect Dictionary* (1964), which in 1998 existed in a handwritten form (about 2000 pages) three times longer than the original version, but he was not able to complete its publication before his death in 2002.

See also: Agrell, Sigurd (1881–1937).

Bibliography

Jarring G (1933). *Studien zu einer osttürkischen Lautlehre.* Lund: Borelius.

Jarring G (1939). *On the distribution of Turk tribes in Afghanistan: an attempt at a preliminary classification.* Lund: Gleerup.

Jarring G (1946–1951). *Materials to the knowledge of Eastern Turki: tales, poetry, proverbs, riddles, ethnological and historical texts from the southern parts of Eastern Turkestan.* Lund: Gleerup.

Jarring G (1964). *An Eastern Turki-English dialect dictionary.* Lund: Gleerup.

Jarring G (1981). *Memoarer 1939–1952.* Stockholm: Bonnier.

Jarring G (1985). *The Moen collection of Eastern Turki (New Uighur) proverbs and popular sayings.* Malmö: LiberFörlag/Gleerup.

Jarring G (1991). *Culture clash in central Asia: Islamic views on Chinese theatre: Eastern Turki texts.* Stockholm: Almqvist & Wiksell International.

Jarring G (1991). *Prints from Kashgar: the printing-office of the Swedish Mission in Eastern Turkestan: history and production with an attempt at a bibliography.* Stockholm: Almqvist & Wiksell International.

Jarring G (1997). *Central Asian Turkic place-names: Lop Nor and Tarim area: an attempt at classification and explanation based on Sven Hedin's diaries and published works.* Stockholm: Sven Hedin Foundation, National Museum of Ethnography.

Rosén S (2003). 'Gunnar Jarring.' In Kungl *Vitterhets Historie och Antikvitets Akademiens Årsbok 2003.* Stockholm: Almqvist & Wiksell International. 34–40.

Javanese

B Nothofer, Universität Frankfurt, Frankfurt, Germany

Javanese is most-spoken regional language of Indonesia and most-spoken language of the Austronesian language family with about 75 million speakers. It is spoken along the northwest coast of Java (Banten, Krawang, Cirebon) and in the central and eastern areas of this island (68 million speakers). Outside of Java, it is used in the Indonesian transmigration areas of Sumatra, Kalimantan, and Sulawesi (altogether 8 million speakers) as well as in Suriname (60 000 speakers) and in New Caledonia (6700 speakers). Of the three dialects usually distinguished (western, central, and eastern) (Ras, 1994), the western one was

ethnological and historical texts from the southern parts of Eastern Turkestan (1946–1991) and *The Moen collection of Eastern Turki (New Uighur) popular poetry* (1996). He treated the demographic situation in *On the distribution of Turk tribes in Afghanistan: an attempt at a preliminary classification* (1939). As an example of his religious observations one may mention *Culture clash in central Asia: Islamic views on Chinese theatre: Eastern Turki texts* (1991). During his stay as ambassador in Moscow from 1964 to1973, Jarring became interested in the history of Swedish–Russian relations, including the interesting destinies of the prisoners from the wars of Charles XII (the Caroleans). Jarring also examined the place-names noted by his famous compatriot Sven Hedin and published *Central Asian Turkic place-names: Lop Nor and Tarim area: an attempt at classification and explanation based on Sven Hedin's diaries* (1997).

Jarring became engaged in Swedish diplomatic service in 1941 and served in Teheran, New Delhi, Washington, and Moscow. He was the Swedish representative in the United Nations (UN) security council in 1957–1958. He was the UN secretary's special representative for the Middle East in 1967–1990 and served as mediator in the Kashmir conflict and in the conflict after the Six-Day War between Egypt and Israel. His name was also suggested for the post of secretary of the United Nations. He has reported about his life and experience in several books, e.g., *Memoarer 1939–1952* (1981).

As an international diplomat Jarring was known as 'The Clam,' and it was said about him that 'few could keep silent in as many languages as he.' In spite of his many important diplomatic tasks he managed to work continually in his special linguistic and ethnographic fields. His list of publications includes over 1000 items. He continued to extend his *Eastern Turki-English Dialect Dictionary* (1964), which in 1998 existed in a handwritten form (about 2000 pages) three times longer than the original version, but he was not able to complete its publication before his death in 2002.

See also: Agrell, Sigurd (1881–1937).

Bibliography

Jarring G (1933). *Studien zu einer osttürkischen Lautlehre.* Lund: Borelius.

Jarring G (1939). *On the distribution of Turk tribes in Afghanistan: an attempt at a preliminary classification.* Lund: Gleerup.

Jarring G (1946–1951). *Materials to the knowledge of Eastern Turki: tales, poetry, proverbs, riddles, ethnological and historical texts from the southern parts of Eastern Turkestan.* Lund: Gleerup.

Jarring G (1964). *An Eastern Turki-English dialect dictionary.* Lund: Gleerup.

Jarring G (1981). *Memoarer 1939–1952.* Stockholm: Bonnier.

Jarring G (1985). *The Moen collection of Eastern Turki (New Uighur) proverbs and popular sayings.* Malmö: LiberFörlag/Gleerup.

Jarring G (1991). *Culture clash in central Asia: Islamic views on Chinese theatre: Eastern Turki texts.* Stockholm: Almqvist & Wiksell International.

Jarring G (1991). *Prints from Kashgar: the printing-office of the Swedish Mission in Eastern Turkestan: history and production with an attempt at a bibliography.* Stockholm: Almqvist & Wiksell International.

Jarring G (1997). *Central Asian Turkic place-names: Lop Nor and Tarim area: an attempt at classification and explanation based on Sven Hedin's diaries and published works.* Stockholm: Sven Hedin Foundation, National Museum of Ethnography.

Rosén S (2003). 'Gunnar Jarring.' In Kungl *Vitterhets Historie och Antikvitets Akademiens Årsbok 2003.* Stockholm: Almqvist & Wiksell International. 34–40.

Javanese

B Nothofer, Universität Frankfurt, Frankfurt, Germany

Javanese is most-spoken regional language of Indonesia and most-spoken language of the Austronesian language family with about 75 million speakers. It is spoken along the northwest coast of Java (Banten, Krawang, Cirebon) and in the central and eastern areas of this island (68 million speakers). Outside of Java, it is used in the Indonesian transmigration areas of Sumatra, Kalimantan, and Sulawesi (altogether 8 million speakers) as well as in Suriname (60 000 speakers) and in New Caledonia (6700 speakers). Of the three dialects usually distinguished (western, central, and eastern) (Ras, 1994), the western one was

Table 1 Javanese consonants

Consonant type	*Labial*	*Dental*	*Retroflex*	*Palatal*	*Velar*	*Glottal*
Voiceless stops	*p*	*t*	*ṭ*	*c*	*k*	*ʔ*
'Voiced' stops	*b*	*d*	*ḍ*	*j*	*g*	
Nasal	*m*		*n*	*ɲ*	*ŋ*	
Fricative			*s*			*h*
Approximant	*w*		*r, l*	*y*		

divided into seven subdialects (Nothofer, 1980). The central and eastern dialect variants have not yet been studied in much detail.

Standard Javanese is the language as it is spoken in the area of Surakarta and Yogyakarta. Javanese has speech levels that are based on the principle to whom and about whom one talks. The level chosen depends on factors such as age, status, and respect. The degree of politeness is expressed by lexical or affixal choices. The speech levels are a Javanese innovation. Clynes (1992) argues that this system was well established by the 15th century. Sundanese (Sunda), Madurese (Madura), Balinese (Bali), and Sasak (but not Malay) have borrowed these speech levels. The position of Javanese in the Western Malayo-Polynesian subfamily has been a matter of dispute. While Dyen (1965) and Nothofer (1975) grouped Javanese with Malay, Sundanese, and Madurese, Nothofer (1985) suggests that the latter three – although not necessarily constituting a subgroup – are more closely related to each other than they are to Javanese. Javanese is one of the few Austronesian languages whose history can be traced because of the existence of older texts. The oldest records date back to the 8th century. Zoetmulder (1974) deals with Old Javanese literature and Zoetmulder (1982) is an Old Javanese-English dictionary.

Javanese Phonology

Consonants

The Javanese consonant system resembles that of other languages of western Indonesia. There are features that are common to Javanese and Madurese only. Both languages share a phonemic distinction between dental and retroflex stops. The 'voiced' consonants of Javanese are pronounced like voiceless stops with breathy voice of the following vowel (Fagan, 1988; Arps *et al.*, 2000). The consonants are shown in Table 1.

Vowels

Nothofer (1980) suggests a system of six vowel phonemes, shown in Table 2.

Javanese has the following allophonic rules (Clynes, 1995; Nothofer, 1980): /*i*/, /*u*/ are realized as [*I*], [*ʊ*] in closed syllables and as [*i*], [*u*] elsewhere. /*ə*/ is always realized as [ə]. The phonemes /*e*/, /*o*/ are realized as [*ɛ*] and [*ɔ*] in closed syllables, in open syllables where the vowel in a following open syllable is high, and in open syllables where the vowel in a following syllable is identical or /*ə*/. In all other positions, these phonemes are realized as [*e*] and [*o*]. The phoneme /a/ appears as [*ɔ*] word-finally and in penultimate open syllables where a following open syllable has /a/. Otherwise, it appears as [*a*]. Allophonic variation also depends on the initial phoneme of suffixes: the addition of a consonant-initial suffix results in the treatment of the stem-final vowel as if it appeared in a closed syllable and the addition of a vowel-initial suffix will cause a high vowel in the stem-final closed syllable to behave as if it appeared in an open syllable.

Morphology

Verbal affixes include the following: *N-* indicates an 'active' and *di-* a 'passive' transitive verb, *kə-* marks 'accidental passive.' *-ke* forms transitive verbs whose patient is the causee or benefactee, while *-i* forms transitive verbs in which the undergoer is the location or goal of the action. Intransitive verbs with *-an* indicate 'nonchalance, to be in the state of, to busy o.s. with.' The suffix *-a* with the sense 'in case, assuming, although' occurs with nonverbs (e.g., pronouns, adjectives) and verbs. Arps *et al.* (2000) call it the 'irreality-suffix.' It is distinguished from the verbal suffix *-a,* which is an imperative marker. The suffix *-ən* is added to nouns denoting physiological conditions and forms verbs meaning 'undergo the physiological process of (the noun).' The nominal affixes include *-an,* which derives nouns that stand as objects of an action indicated by the verb; *kə- -an*, used to nominalize qualities; nouns formed with *pəN-* refer to a person carrying out the action of the verb or an instrument with which the action is performed; and *pə-an* nouns refer to a process or result of an action indicated by the verb or to the location where an action of the corresponding verb occurs.

The meaning of total reduplication of nouns is 'diversity, completeness,' while that of verbs has

Table 2 Javanese vowel phonemes

Vowel type	*Front*	*Central*	*Back*
High	*i*		*u*
Mid	*e*	*ə*	*o*
Low		*a*	

the sense of 'durative, intensive, iterative.' Verbal reduplication can also involve vowel variation in the first member of the doubled form.

Writing System

After the World War II, the publication of texts in Javanese script (hanacaraka or caraka) came to an end. Actually, the Latin script began to replace the Javanese alphabet at the beginning of the 20th century. The traditional script originates from a Pallava script of southern India.

See also: Austronesian Languages: Overview; Indonesia: Language Situation.

Bibliography

Arps B *et al.* (2000). *Semaian 20: Hedendaags Javaans. Een leerboek*. Leiden: Opleiding Talen en Culturen van Zuidoost-Azië en Oceanië, Universiteit Leiden.

Clynes A (1992). 'Old Javanese influence on Balinese: speech styles.' In Dutton T & Tryon D T (eds.) *Language contact and change in the Austronesian world*. Berlin, New York: Mouton de Gruyter.

Clynes A (1995). 'Javanese.' In Tryon D T (ed.) *An introduction to Austronesian studies 1: Fascicle 1: Comparative Austronesian dictionary.* Berlin, New York: Mouton de Gruyter. 469–484.

Dyen I (1965). *A lexicostatistical classification of the Austronesian languages. IUPAL memoir 19; supplement to IJAL 25*. Baltimore: Waverly Press.

Errington J J (1998). *Studies in the Social and Cultural Foundations of Language 19: Shifting languages: Interaction and identity in Javanese Indonesia*. Cambridge: Cambridge University Press.

Fagan J L (1988). 'Javanese intervocalic stop phonemes: the light/heavy distinction.' In McGinn R (ed.) *Studies in Austronesian linguistics*. Athens, Ohio: Center for Southeast Asia Studies, Ohio University.

Keeler W (1984). *International Studies, Southeast Asian Series 69: Javanese: A cultural approach*. Athens: Ohio University Center for International Studies.

Molen W van der (1993). *Semaian 8: Javaans schrift.* Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Rijksuniversiteit Leiden.

Nothofer B (1975). *Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 73: The reconstruction of Proto-Malayo-Javanic.* The Hague: Nijhoff.

Nothofer B (1980). *Dialektgeographische Untersuchungen in West-Java und im westlichen Zentral-Java.* Wiesbaden: Harrassowitz.

Nothofer B (1981). *Dialektatlas von Zentral-Java.* Wiesbaden: Harrassowitz.

Nothofer B (1985). 'The subgrouping of the languages in the Javo-Sumatra Hesion: a reconsideration.' *Bijdragen tot de Taal-, Land- en Volkenkunde 141(2/3),* 288–302.

Pigeaud Th (n. d.). *Nederlands-Javaans handwoordenboek*. Groningen: Wolters-Noordhoff.

Ras J J (1994). *Inleiding tot het modern Javaans.* Leiden: KITLV Uitgeverij.

Sudaryanto & Pranowo (eds.) (2001). *Kamus Pepak Basa Jawa*. Yogyakarta: Badan Pekerja Kongres Bahasa Jawa.

Sudaryanto *et al.* (1991). *Tata bahasa baku bahasa-Jawa.* Yogyakarta: Duta Wacana University Press.

Zoetmulder P J (1974). *Kalangwan: a survey of Old Javanese literature*. The Hague: Nijhoff.

Zoetmulder P J (1982). *Old Javanese dictionary* (2 vols). The Hague: Nijhoff.

Jèrriais

M C Jones, University of Cambridge, Cambridge, UK

Background

Jèrriais is the dialect spoken on Jersey, the largest of the Channel Islands. It is related to the Norman dialects of northern France. (For details of Jersey's linguistic history, including its Anglicization and the decline in speaker numbers, *see* **Channel Islands: Language Situation**.)

According to the 2001 Census of Jersey, there then remained only 2874 speakers of Jèrriais, (3.2% of the total resident population). Some two-thirds of these were over age 60 and only 113 speakers declared Jèrriais to be their usual everyday language.

Phonological Structure

Although the phonological structure of Jèrriais is similar to that of standard French, there are also significant differences. Marked regional variation is still very much in evidence in modern Jèrriais, with

Table 2 Javanese vowel phonemes

Vowel type	*Front*	*Central*	*Back*
High	*i*		*u*
Mid	*e*	*ə*	*o*
Low		*a*	

the sense of 'durative, intensive, iterative.' Verbal reduplication can also involve vowel variation in the first member of the doubled form.

Writing System

After the World War II, the publication of texts in Javanese script (hanacaraka or caraka) came to an end. Actually, the Latin script began to replace the Javanese alphabet at the beginning of the 20th century. The traditional script originates from a Pallava script of southern India.

See also: Austronesian Languages: Overview; Indonesia: Language Situation.

Bibliography

Arps B *et al.* (2000). *Semaian 20: Hedendaags Javaans. Een leerboek*. Leiden: Opleiding Talen en Culturen van Zuidoost-Azië en Oceanië, Universiteit Leiden.

Clynes A (1992). 'Old Javanese influence on Balinese: speech styles.' In Dutton T & Tryon D T (eds.) *Language contact and change in the Austronesian world*. Berlin, New York: Mouton de Gruyter.

Clynes A (1995). 'Javanese.' In Tryon D T (ed.) *An introduction to Austronesian studies 1: Fascicle 1: Comparative Austronesian dictionary.* Berlin, New York: Mouton de Gruyter. 469–484.

Dyen I (1965). *A lexicostatistical classification of the Austronesian languages. IUPAL memoir 19; supplement to IJAL 25*. Baltimore: Waverly Press.

Errington J J (1998). *Studies in the Social and Cultural Foundations of Language 19: Shifting languages: Interaction and identity in Javanese Indonesia*. Cambridge: Cambridge University Press.

Fagan J L (1988). 'Javanese intervocalic stop phonemes: the light/heavy distinction.' In McGinn R (ed.) *Studies in Austronesian linguistics*. Athens, Ohio: Center for Southeast Asia Studies, Ohio University.

Keeler W (1984). *International Studies, Southeast Asian Series 69: Javanese: A cultural approach*. Athens: Ohio University Center for International Studies.

Molen W van der (1993). *Semaian 8: Javaans schrift*. Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Rijksuniversiteit Leiden.

Nothofer B (1975). *Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde 73: The reconstruction of Proto-Malayo-Javanic*. The Hague: Nijhoff.

Nothofer B (1980). *Dialektgeographische Untersuchungen in West-Java und im westlichen Zentral-Java*. Wiesbaden: Harrassowitz.

Nothofer B (1981). *Dialektatlas von Zentral-Java*. Wiesbaden: Harrassowitz.

Nothofer B (1985). 'The subgrouping of the languages in the Javo-Sumatra Hesion: a reconsideration.' *Bijdragen tot de Taal-, Land- en Volkenkunde 141(2/3),* 288–302.

Pigeaud Th (n. d.). *Nederlands-Javaans handwoordenboek*. Groningen: Wolters-Noordhoff.

Ras J J (1994). *Inleiding tot het modern Javaans*. Leiden: KITLV Uitgeverij.

Sudaryanto & Pranowo (eds.) (2001). *Kamus Pepak Basa Jawa*. Yogyakarta: Badan Pekerja Kongres Bahasa Jawa.

Sudaryanto *et al.* (1991). *Tata bahasa baku bahasa-Jawa*. Yogyakarta: Duta Wacana University Press.

Zoetmulder P J (1974). *Kalangwan: a survey of Old Javanese literature*. The Hague: Nijhoff.

Zoetmulder P J (1982). *Old Javanese dictionary* (2 vols). The Hague: Nijhoff.

Jèrriais

M C Jones, University of Cambridge, Cambridge, UK

Background

Jèrriais is the dialect spoken on Jersey, the largest of the Channel Islands. It is related to the Norman dialects of northern France. (For details of Jersey's linguistic history, including its Anglicization and the decline in speaker numbers, *see* **Channel Islands: Language Situation**.)

According to the 2001 Census of Jersey, there then remained only 2874 speakers of Jèrriais, (3.2% of the total resident population). Some two-thirds of these were over age 60 and only 113 speakers declared Jèrriais to be their usual everyday language.

Phonological Structure

Although the phonological structure of Jèrriais is similar to that of standard French, there are also significant differences. Marked regional variation is still very much in evidence in modern Jèrriais, with

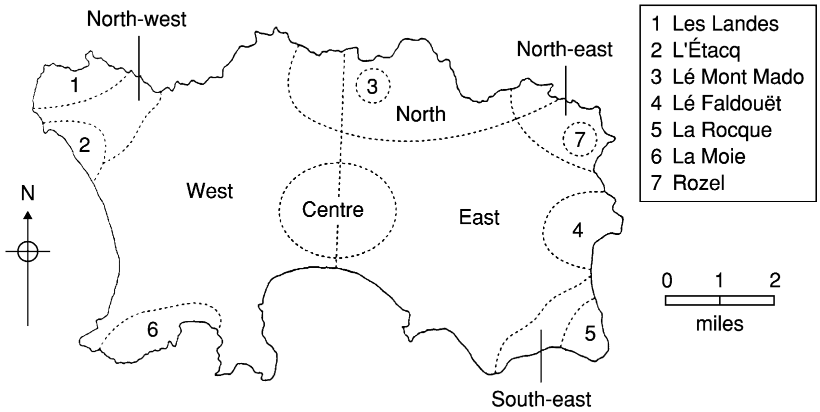


Figure 1 Jersey's 'linguistic pockets.' Reproduced with permission from Jones M C (2001). *Jersey Norman French: a linguistic study of an obsolescent dialect*. Oxford: Blackwell.

the sub-varieties usually categorized in two main groups: East and West Jèrriais (see Figure 1). Even more localized variation is readily observable, although with the decline in speaker-numbers, these so-called linguistic pockets are fast disappearing. This internal variation has never been based on any administrative or other territorial boundaries within Jersey, but many of the Islanders feel it to be intrinsically linked with parish boundaries (see Figure 2), and the practice of using the name of parishes to refer to the sub-dialects of Jèrriais is well established, although not strictly correct. Unless stated, the following comments relate to most varieties of Jèrriais.

Vowels:
Oral: /i e ɛ y œ a u o ə/; the French vowels /ø/ and /ɔ/ are lacking
Nasal: /ẽ ɛ̃ ø̃ õ ɑ̃/;

Unlike in standard French, vowel length is phonemic and all vowels, except /ə/, can be either short or long. Long /a/ is usually realized phonetically as [ɑ]. Jèrriais therefore has 27 vowel phonemes (17 oral, 10 nasal) compared to the 16 of standard French.

Consonants:
Stop: /p b t d k g/
Fricative: /f v s z ð ʃ ʒ h/
Affricate: /tʃ dʒ/
Nasal: /m n ɲ/
Lateral: /l/
Trilled: /r/; this corresponds to the uvular *r* of standard French

/ð/ (written *th*) occurs mainly as a result of the assibilation of intervocalic /r/ in western Jersey and some northern parts of Trinité and St. Martin e.g., *dithe* 'to say' (Fr. *dire*). In St. Ouen, /ð/ also occurs as a development of intervocalic /z/, e.g., *maîson* /meðõ/ 'house' (/me(j)zõ/ elsewhere).

In standard French, the affricates /tʃ/ and /dʒ/ only occur in borrowings, e.g., *match*, *gin*. However, in Jèrriais they can occur as the result of the secondary palatalization of /k/ and /g/, e.g., *tchoeu* /tʃœ/ 'heart', *dgèrre* /dʒɛ:r/ 'war' (Fr. *coeur*, *guerre*) and of /t/ and /d/, e.g., *mêtchi* /metʃi/ 'profession', *dgix* /dʃi/ 'ten' (Fr. *métier*, *dix*).

Like standard French, Jèrriais preserves no trace of Latin /h/. /h/ was introduced into the French phonemic system in words borrowed from the language of the Germanic invaders who dominated northern Gaul in the 5th to the 8th century. The sound disappeared from standard French in the early modern period (16th–18th century) but remains in Jèrriais, e.g., *housse* /hus/ 'holly' (Fr. *houx*).

In many parts of Jersey, when /l/ is the second element of a consonant cluster, it frequently undergoes delateralization to /j/, e.g., *clios* /kjo/ 'field' (Fr. *clos*). In St. Ouen, however, the more conservative form /kʎo/ is still to be heard amongst older speakers. St. Ouen also retains the /ʎ/ pronunciation word-finally in words such as /fiʎ/ *fil'ye* ('daughter') whereas, elsewhere in Jersey, /ʎ/ has become depalatalized to /l/.

The velar nasal /ŋ/ occurs in standard French only in borrowings, e.g., *le shopping* /ʃɔpɪŋ/. Despite its restricted distribution, it is generally given phonemic status. The sound also occurs in Jèrriais in English borrowings such as *blanket*, *dinghy* /blæŋkɛt/, /dɪŋi/. However in Jèrriais /ŋ/ is not considered to be phonemic.

Glides The three glides of standard French (/j ɥ w/) also occur in Jèrriais but there is a tendency for /ɥ/ to be replaced by /w/.

Vocabulary

Most of the vocabulary of Jèrriais is shared with standard French, but regional variation is also

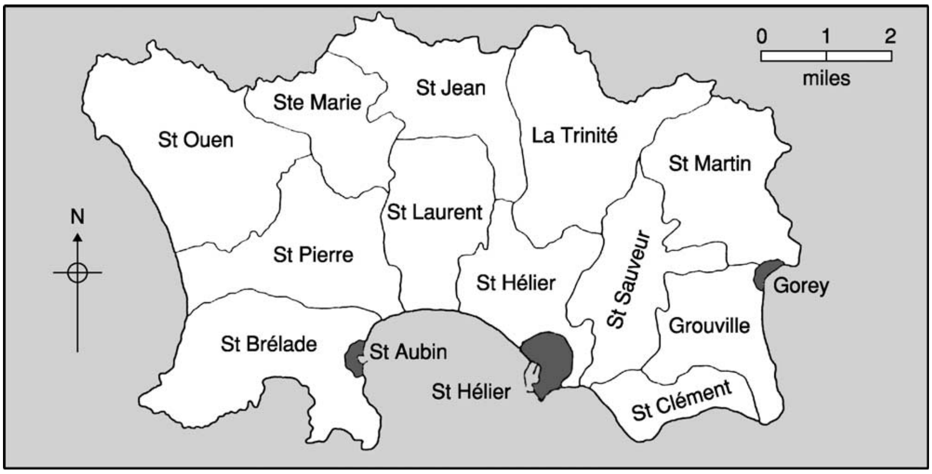


Figure 2 The twelve parishes of Jersey. Reproduced with permission from Jones M C (2001). *Jersey Norman French: a linguistic study of an obsolescent dialect.* Oxford: Blackwell.

apparent in the lexis of Jèrriais, e.g., *pêtre* 'spider' (WJ), *ithangnie* (EJ). The dominant influence of English has had far-reaching linguistic consequences, most noticeably in the lexis, where borrowings abound in many everyday domains. Some of these are well established, e.g., *bouchet* 'bucket', *ticl'ye* '(tea)kettle' while others are more recent, e.g., *software*. The semantic adaptation of Jèrriais words on the basis of their English equivalents is also found, e.g., *J'n'sais pon comment chenna* ***travaille*** 'I don't know how that **works**', as are calques of English phrasal verbs and other expressions, e.g., *i'tchit bas* 'he fell down'; *j'chèrchis pouor* 'I looked for it'.

Morphosyntax

The more striking morphosyntactic differences between Jèrriais and standard French include:

i. the Old French distinction of number in masculine nouns and adjectives between *–el* (s.), *-eaus* (pl.), e.g., *chastel* 'castle', pl. *chasteaus*; *novel* 'new', pl. *noveaus* 'new', has been lost in standard French as a result of the creation of a new analogical singular based on the plural, e.g., *château* ~ *châteaux* (both pronounced [ʃato]), *nouveau* ~ *nouveaux*, though the original singular remains before vowels, e.g., *le nouvel an* 'New Year'; in Jèrriais, the distinction is maintained, *châté* ~ *châtchieaux*, *nouvé* ~ *nouvieaux*;
ii. the first person plural personal pronoun subject *nous* is replaced by *jé/j'*, e.g., *j'pâlons* (Fr. *nous parlons*) 'we speak';
iii. there is no specific feminine third person plural subject personal pronoun, *i'* serving for both genders (cf. Fr. *ils* masc., *elles* fem.);
iv. adjectives of color almost invariably precede the noun, e.g., *lé nièr cat* 'the black cat' (Fr. *le chat noir*);
v. the preterite tense, which has gone out of use in informal spoken French, though it survives in the formal written language, is widely used in spoken and written Jèrriais, e.g., *j'donnis* 'I gave'. Certain third person plural preterite forms are restricted to St. Ouen, e.g., *i'vidrent* 'they saw', *i'fûdrent* 'they went' (for *i'vîtent*, *i'fûtent* elsewhere)
vi. the imperfect subjunctive, now virtually defunct in spoken French, also survives, e.g., *j'voulais qu'i'l'sûsse* 'I wanted him to know'.

The dominance of English on Jersey is also leading to an increase in frequency of syntactic constructions more isomorphic with English.

Language Planning

Since being introduced into the education system in 1999, Jèrriais is now offered on an extra-curricular basis in most of the Island's schools. It also features in a couple of weekly radio slots and in a fortnightly newspaper column. However, Jèrriais enjoys no more than a token presence on television. Language planning measures receive very little official backing, and with the exception of the Jèrriais education initiative, which has received funding from the States of Jersey, have been left in the hands of groups of enthusiasts.

Jèrriais has been codified via a dictionary (Le Maistre, 1966) and grammar (Birt, 1985). The standard variety is based largely on the sub-dialect of St. Ouen.

Literary Tradition

Although the important 12th-century writer Wace (c. 1100–1179) is known to have come from Jersey,

no literary writings in Jèrriais exist until the 19th century. The first author to use the dialect as a medium for his work was Matthieu Le Geyt (1777–1849). Jersey produced several poets and writers during the course of the 19th and 20th centuries; much of their output was published in newspapers and periodicals. Although collections of satirical short stories in Jèrriais were published in pamphlet form from the third quarter of the 19th century onward, the first complete volume of prose to be published in Jèrriais was George Le Feuvre's *Jèrri jadis* (1973).

See also: Channel Islands: Language Situation; French.

Bibliography

Birt P (1985). *Lé Jèrriais pour tous. A complete course on the Jersey language.* Jersey: Don Balleine.

Brasseur P (1978a). 'Les principales caractéristiques phonétiques des parlers normands de Jersey, Sercq, Guernesey et Magneville (canton de Bricquebec, Manche), Première partie.' *Annales de Normandie 25(1),* 49–64.

Brasseur P (1978b). 'Les principales caractéristiques phonétiques des parlers normands de Jersey, Sercq, Guernesey et Magneville (canton de Bricquebec, Manche), Deuxième partie.' *Annales de Normandie 25(3),* 275–306.

Carré A L (1972). *English-Jersey language vocabulary.* Jersey: Don Balleine.

Falle P (1734). *An account of the island of Jersey* (2nd edn.). London: John Newton.

Hublart C (1979). 'Le français de Jersey.' Unpublished thesis, Université de l'Etat à Mons.

Jones M C (2000a). 'Swimming against the tide: language planning on Jersey.' *Language Problems and Language Planning 24(2),* 167–196.

Jones M C (2000b). 'Ambiguity and unpredictability: linguistic change in modern Jèrriais.' *Verbum 22(2),* 203–222.

Jones M C (2001). *Jersey Norman French: a linguistic study of an obsolescent dialect.* Oxford: Blackwell.

Jones M C (2005a). 'Transfer and changing linguistic norms in Jersey Norman French.' *Bilingualism: Language and Cognition 8(2)*, 1–17.

Jones M C (2005b). 'Some structural and social correlates of single word intrasentential codeswitching in Jersey Norman French.' *Journal of French Language Studies 15 (1)*, 1–23.

Joret C (1883). *Des caractères et de l'extension du patois normand.* Paris: Vieweg.

Lebarbenchon R J (n.d.). *La Grève de Lecq. Littératures et cultures populaires de Normandie 1: Guernesey et Jersey.* Cherbourg: Isoète.

Le Feuvre G F (1973). *Jèrri jadis.* Jersey: Don Balleine. [2nd edn. 1983.]

Le Feuvre G F (1976). *Histouaithes et gens d'Jèrri.* Jersey: Don Balleine.

Le Maistre F (1947). 'The Jersey language in its present state. The passing of a Norman heritage.' Paper read before the Jersey Society in London (July 1947). London: The Jersey Society.

Le Maistre F (1966). *Dictionnaire Jersiais-Français.* Jersey: Don Balleine.

Le Maistre F (1979) (1993). *The Jersey language.* (5 cassettes/booklets). Jersey: Don Balleine.

Liddicoat A J (1994). *A grammar of the Norman French of the Channel Islands: the dialects of Jersey and Sark.* Berlin: Mouton de Gruyter.

Mason I (1980). 'L'usage des temps narratifs en jersiaise et val de sairais.' *Parlers et Traditions Populaires de Normandie 47–48,* 99–106.

Scott Warren A (1999). *Lé neu c'mîn.* Jersey: Don Balleine.

Société Jersiaise (2000). *Les preunmié mille mots.* Jersey: Société Jersiaise.

Spence N C W (1957a). 'Jèrriais and the dialects of the Norman mainland.' *Bulletin de la Société Jersiaise 17,* 81–90.

Spence N C W (1957b). 'L'Assibilation de l'r intervocalique dans les parlers jersiais.' *Revue de Linguistique Romane 21,* 270–288.

Spence N C W (1960). *Glossary of Jersey French.* Oxford: Blackwell.

Spence N C W (1985). 'Phonologie descriptive des parlers jersiais: I. Les voyelles.' *Revue de Linguistique Romane 49,* 151–165.

Spence N C W (1987). 'Phonologie descriptive des parlers jersiais: II. Les consonnes.' *Revue de Linguistique Romane 51,* 119–133.

Spence N C W (1993). *A brief history of Jèrriais.* Jersey: Don Balleine.

Syvret M & Stevens J (1998). *Balleine's history of Jersey.* West Sussex: Phillimore & Co.

Tapley J (1998). *Jersey Norman French is fun.* Jersey: Don Balleine.

Jespersen, Otto (1860–1943)

M Pierce, University of Michigan, Ann Arbor, MI, USA

Otto Jespersen was born on July 16, 1860, in Randers, Denmark. In 1870 his father died, and 4 years later, his mother died as well. Although Jespersen had been very interested in languages as a child and had studied a number of them on his own (including Italian, Spanish, Greek, and Old Norse), following family tradition, he originally studied law at the University of Copenhagen, where he enrolled in 1877. After 4 apparently unhappy years of studying law, during which time he partially supported himself by working as a shorthand reporter in the Danish Parliament, he switched fields to Romance languages. In 1887 he was awarded a master's degree in French and spent the next year abroad, in England, France, and Germany, improving his knowledge of the local languages, establishing contacts with various scholars, and studying phonetics and Old and Middle English. On his return to Denmark, Jespersen began teaching French and English at various schools. In 1891 he was awarded a doctorate from the University of Copenhagen for a dissertation about the case system of English. He continued school teaching until 1893, when he became professor of English at the University of Copenhagen, where he would remain (excluding visiting positions) for the remainder of his career. Jespersen served as visiting professor at Columbia University and the University of California at Berkeley in 1909–1910 and as rector of the University of Copenhagen in 1920–1921. He retired from the University of Copenhagen in 1925 but continued to publish at a prodigious rate until his death in 1943.

Jespersen was the recipient of numerous scholarly honors, receiving honorary doctorates from Columbia University (1910), St. Andrews University (1925), and the Sorbonne (1927). He was invited to lecture at the Congress of Arts and Sciences, held in conjunction with the World Fair, in St. Louis in 1904. He was elected to the Danish Videnskabernes Skelskab in 1899 and served as president of the Fourth International Congress of Linguists, held in Copenhagen in 1936. He had a wide circle of students, friends, and admirers, as evidenced by the two separate volumes of papers dedicated to him.

Jespersen was an exceptionally productive scholar who contributed to a number of fields within linguistics. He is perhaps best known for his work on English, phonetics, general linguistics, language pedagogy, and historical linguistics. His most important contribution to English linguistics is certainly his monumental *Modern English grammar on historical principles* (published in seven parts, 1909–1949). Although the organization of the work may seem random at first glance, as volume 1 deals with the history of English phonology, volumes 2–5 with syntax, and volume 6 with morphology, and then volume 7 returns to syntax, it is one of the truly great books of linguistics and a fitting culmination of a life's work on linguistics and the English language. Jespersen also wrote a number of shorter works on English linguistics, including his outstanding *Growth and structure of the English language* (Jespersen, 1938), which was used for many years in university classes and translated into Japanese. The quality of this work is perhaps best indicated by the fact that it won the Volney Prize from the French Institute (Juul *et al*., 1995: 137).

Jespersen was an outstanding phonetician. Although he seems to have originally become interested in phonetics because of its possible applications in language teaching, he quickly became fascinated with it in its own right. He published a general handbook of phonetics early in his scholarly career (Jespersen, 1897; the work was later translated into German, but unfortunately not into English) and even developed his own system of phonetic transcription. True to what originally piqued his interest, Jespersen did not neglect the role of phonetics in language pedagogy. For example, he prepared a textbook of French in which French words were given in phonetic transcription, with an eye toward improving students' pronunciation from the beginning of their studies.

Yet Jespersen was not limited in his research; he often worked on questions of general linguistics. Perhaps his best known book in this area was *Language, its nature, development, and origin* (Jespersen, 1922), which deals with topics like child language, causes of language change, and the history of linguistics. Jespersen was also deeply concerned with problems of language pedagogy. He prepared textbooks or readers for Danish, French, and English, as well as more general works on the subject. Finally, he also published extensively on historical linguistics, including a groundbreaking paper originally published in 1886, which contested the Neogrammarian dogma that sound laws were exceptionless, and also often treated historical questions in his more general works. His most important contribution to this field is probably his work on efficiency as a cause of language change (see especially Jespersen, 1941). Although it would be misguided to argue that Jespersen saw efficiency as the cause of all language change, he certainly contended that it played an important role. In addition, Jespersen worked on a number of

other subjects not discussed here, including international languages, English literature, and syntax (see Jespersen 1893, 1924, and 1928, among others).

See also: Artificial Languages; English in the Present Day (since ca. 1900); English, Later Modern (ca. 1700–1900); Neogrammarians; Phonetics: Overview.

Bibliography

Jespersen O (1886). 'Zur Lautgesetzfrage.' *Internationale Zeitschrift für allgemeine Sprachwissenschaft 3,* 188–216.

Jespersen O (1893). *Chaucers liv og digtning.* Copenhagen: Klein.

Jespersen O (1897). *Fonetik: En systematisk fremstilling af loeren om sproglyd.* Copenhagen: Schubothe.

Jespersen O (1909–1949). *Modern English grammar on historical principles.* Copenhagen: Munksgaard.

Jespersen O (1922). *Language: its nature, development, and origin.* New York: Holt.

Jespersen O (1924). *The philosophy of grammar.* London: Allen and Unwin.

Jespersen O (1928). *An international language.* London: Allen and Unwin.

Jespersen O (1938). *Growth and structure of the English language* (9th edn.). Oxford: Blackwell.

Jespersen O (1941). *Efficiency in language change.* Copenhagen: Munksgaard.

Juul A & Nielsen H F (eds.) (1989). *Otto Jespersen: Facets of his Life and Work.* Amsterdam: Benjamins.

Juul A, Nielsen H F & Nielsen J E (eds. and trans.) (1995). *A linguist's life. An English translation of Otto Jespersen's autobiography with notes, photos, and a bibliography.* Odense: Odense University Press.

Jewish Languages

B Spolsky, Bar-Ilan University, Ramat Gan, Israel
S B Benor, Hebrew Union College – Jewish Institute of Religion, Los Angeles, CA, USA

Introduction

Ethnic and religious groups use language as one means of constructing and expressing their distinctness from other groups. Jews are no exception. Wherever Jews have lived – from Baghdad to Brooklyn, Amsterdam to Odessa – they have spoken somewhat differently from their non-Jewish neighbors. These differences have been as small as the addition of a few Hebrew words and as large as a vastly different lexicon, syntax, and phonology. Therefore, the term "Jewish language" refers to any linguistic variety spoken by Jews that differs to some extent from the non-Jewish language(s) around it. The field of Jewish language studies examines the distinct linguistic practices of the Jewish people around the world.

Jewish languages generally exist in a situation of triglossia with the local non-Jewish language(s) and with a liturgical combination of Hebrew and Aramaic (Weinreich, 1980; Rabin, 1981; and Fishman, 1985). The Jewish language is used mostly for intra-community speech and sometimes for writing. Speakers also generally have at least some knowledge of the co-territorial non-Jewish languages and use them in their interactions with non-Jews. Hebrew and Aramaic have played a very important role in Jewish life. Biblical and rabbinic literatures are studied regularly in their original languages, and daily prayers are conducted mostly in Hebrew and Aramaic. Hebrew is also used for contemporary rabbinic and liturgical production, as well as some other literary functions.

Jewish languages have been documented in many parts of the Jewish diaspora: Yiddish (sometimes referred to as Judeo-German), Judeo-Spanish (also called Ladino, Judezmo, Dzhudezmo, Jidyó, Spanyol, Spanyolit), Judeo-Greek (Yevanic, Romaniyot), Judeo-Italian (Italkic, including local varieties like Judeo-Venetian), Judeo-Portuguese, Judeo-French (Zarphatic, Western Loez), and Judeo-Provençal (Shuadit) in Europe; Judeo-Arabic (Yahudic), Judeo-Aramaic (Targum, Kurdit), Judeo-Persian (Jidi, Parsic, Judeo-Tadjik, Judeo-Tat, Bukharan), Judeo-Georgian (Gurjuc, Gruzinic), Judeo-Crimean Tatar (Krimchak), and Judeo-Berber in the Middle East, North Africa, and the former Soviet Union; Judeo-Malayalam in India; and Jewish English (Yinglish, Yeshivish) in the New World. Some of the larger and better-studied cases are described in separate sections below.

Scholarship

The scholarly recognition of a phenomenon of Jewish languages goes back to the beginning of the 20th century, when Yiddish became the object of serious academic study. Mieses (1915) presented the first large-scale exploration of Jewish linguistic varieties. In late 1970s and early 1980s, Jewish languages

other subjects not discussed here, including international languages, English literature, and syntax (see Jespersen 1893, 1924, and 1928, among others).

See also: Artificial Languages; English in the Present Day (since ca. 1900); English, Later Modern (ca. 1700–1900); Neogrammarians; Phonetics: Overview.

Bibliography

Jespersen O (1886). 'Zur Lautgesetzfrage.' *Internationale Zeitschrift für allgemeine Sprachwissenschaft 3,* 188–216.

Jespersen O (1893). *Chaucers liv og digtning.* Copenhagen: Klein.

Jespersen O (1897). *Fonetik: En systematisk fremstilling af loeren om sproglyd.* Copenhagen: Schubothe.

Jespersen O (1909–1949). *Modern English grammar on historical principles.* Copenhagen: Munksgaard.

Jespersen O (1922). *Language: its nature, development, and origin.* New York: Holt.

Jespersen O (1924). *The philosophy of grammar.* London: Allen and Unwin.

Jespersen O (1928). *An international language.* London: Allen and Unwin.

Jespersen O (1938). *Growth and structure of the English language* (9th edn.). Oxford: Blackwell.

Jespersen O (1941). *Efficiency in language change.* Copenhagen: Munksgaard.

Juul A & Nielsen H F (eds.) (1989). *Otto Jespersen: Facets of his Life and Work.* Amsterdam: Benjamins.

Juul A, Nielsen H F & Nielsen J E (eds. and trans.) (1995). *A linguist's life. An English translation of Otto Jespersen's autobiography with notes, photos, and a bibliography.* Odense: Odense University Press.

Jewish Languages

B Spolsky, Bar-Ilan University, Ramat Gan, Israel
S B Benor, Hebrew Union College – Jewish Institute of Religion, Los Angeles, CA, USA

Introduction

Ethnic and religious groups use language as one means of constructing and expressing their distinctness from other groups. Jews are no exception. Wherever Jews have lived – from Baghdad to Brooklyn, Amsterdam to Odessa – they have spoken somewhat differently from their non-Jewish neighbors. These differences have been as small as the addition of a few Hebrew words and as large as a vastly different lexicon, syntax, and phonology. Therefore, the term "Jewish language" refers to any linguistic variety spoken by Jews that differs to some extent from the non-Jewish language(s) around it. The field of Jewish language studies examines the distinct linguistic practices of the Jewish people around the world.

Jewish languages generally exist in a situation of triglossia with the local non-Jewish language(s) and with a liturgical combination of Hebrew and Aramaic (Weinreich, 1980; Rabin, 1981; and Fishman, 1985). The Jewish language is used mostly for intra-community speech and sometimes for writing. Speakers also generally have at least some knowledge of the co-territorial non-Jewish languages and use them in their interactions with non-Jews. Hebrew and Aramaic have played a very important role in Jewish life. Biblical and rabbinic literatures are studied regularly in their original languages, and daily prayers are conducted mostly in Hebrew and Aramaic. Hebrew is also used for contemporary rabbinic and liturgical production, as well as some other literary functions.

Jewish languages have been documented in many parts of the Jewish diaspora: Yiddish (sometimes referred to as Judeo-German), Judeo-Spanish (also called Ladino, Judezmo, Dzhudezmo, Jidyó, Spanyol, Spanyolit), Judeo-Greek (Yevanic, Romaniyot), Judeo-Italian (Italkic, including local varieties like Judeo-Venetian), Judeo-Portuguese, Judeo-French (Zarphatic, Western Loez), and Judeo-Provençal (Shuadit) in Europe; Judeo-Arabic (Yahudic), Judeo-Aramaic (Targum, Kurdit), Judeo-Persian (Jidi, Parsic, Judeo-Tadjik, Judeo-Tat, Bukharan), Judeo-Georgian (Gurjuc, Gruzinic), Judeo-Crimean Tatar (Krimchak), and Judeo-Berber in the Middle East, North Africa, and the former Soviet Union; Judeo-Malayalam in India; and Jewish English (Yinglish, Yeshivish) in the New World. Some of the larger and better-studied cases are described in separate sections below.

Scholarship

The scholarly recognition of a phenomenon of Jewish languages goes back to the beginning of the 20th century, when Yiddish became the object of serious academic study. Mieses (1915) presented the first large-scale exploration of Jewish linguistic varieties. In late 1970s and early 1980s, Jewish languages

started to be studied intensively, following the publication of two major studies of Yiddish that discussed them in a historical context (Birnbaum, 1979; Weinreich, 1980). Around this time, scholars in Israel and the United States edited symposia (Rabin *et al.*, 1979), collections of articles (Paper, 1978; Fishman, 1985; Gold, 1989) and a short-lived journal (Gold and Prager, 1981–1987). More recently, there has been a wave of renewed interest in the subject, as evidenced by the Jewish Language Research Website (www.jewish-languages.org) and the Jewish Languages Mailing List (www.jewish-languages. org/ml).

History

The presumed monolingualism of the early kingdoms of Israel and Judah gave way, in the centuries after the Babylonian exile in the 6th century B.C., to a Hebrew-Aramaic bilingualism (Chomsky, 1957). By the end of the Temple period 2000 years ago, these languages were supplemented by a widespread knowledge of Greek, which was used with distinctive Jewish features (Wexler, 1985). Thus, Judeo-Aramaic and Judeo-Greek were the earliest Jewish languages that existed in a diglossic relationship with Hebrew.

Judeo-Aramaic was a Jewish adaptation of the major language of wider communication of the Middle East in the millennium before the Common Era. It grew into an important spoken and written Jewish language in Palestine and in the Jewish Diaspora in Babylon, where it was the main language used in the Babylonian Talmud (Greenfield, 1978; Katz, 1985). Among Jews as well as other inhabitants of the region, it was generally replaced by Arabic as a spoken language as a result of the spread of Islam, but it has continued to the present day as a Jewish language in more isolated regions such as Azerbaijan (Garbell, 1965) and Kurdish Iraq (Sabar, 2002).

The third partner in Palestinian trilingualism was Judeo-Greek, widely adapted in Hellenic colonies in Palestine and used by Diaspora Jews throughout the eastern Mediterranean and later in Italy. Judeo-Greek, also called Yevanic, was replaced in most areas starting in the 4th century. An exception is the communities of Romaniote Jews in Greece, which used Judeo-Greek until the influx of Sephardic Jews in the 16th century, when Judeo-Spanish became the majority language of Jews in Greece. Pockets of Judeo-Greek speakers maintained their language in Ioannina, Chalkida, and elsewhere until they were destroyed in the Nazi Holocaust. Few speakers survive today.

Soon after the Roman destruction of Jewish political independence in Palestine in the second century A.D., Hebrew lost its vitality. But it remained firmly entrenched as the language of Jewish religion and literacy, its transmission supported by a religious educational system. Over the next centuries, Jews in exile picked up local languages and developed their own distinctly Jewish varieties, depending in large measure on the nature of their relations with non-Jewish neighbors. As Jews migrated, they generally lost their former language and adapted linguistically to their new land, incorporating distinctive linguistic features. However, two languages defied this trend: Yiddish and Judeo-Spanish, Jewish varieties of Germanic and Hispanic languages, respectively. These languages continued to be used even centuries after their speakers migrated to new lands, where Slavic and Balkan languages were the norm.

In the modern period, when Jews have been able to integrate more fully into some societies, the distinctness of their languages has generally diminished. Yiddish, Judeo-Spanish, and other Jewish languages with long histories have lost significant numbers of speakers due to the combined effects of the Nazi Holocaust and Jews' cultural and linguistic assimilation into new societies in North America, Europe, and Israel. In the 21st century, Jews generally have full competence in the local languages, vernacular and standard. But their speech also tends to maintain some distinctive features, influenced by Hebrew and Aramaic as well as by the Jewish languages spoken by their ancestors.

Common Linguistic Features

Jewish languages tend to have a number of features in common. Structurally, they are generally based on a spoken variety of a non-Jewish language (Yiddish was based on medieval German and Judeo-Spanish on 15th century Spanish), with a large proportion of borrowings from Hebrew and Aramaic, from earlier Jewish languages, and from other contact languages (Weinreich, 1980). In addition, contemporary Jewish languages tend to be influenced by Israeli Hebrew as a result of affiliations with the State of Israel (Benor, 2004).

The Hebrew and Aramaic influences on Jewish languages are mostly lexical, but some phonological and morphosyntactic influences have been documented as well. Hebrew and Aramaic loan words are most common in the semantic fields of religious life, names of individuals and groups, and euphemism. Until recently, Jewish languages were generally written in Hebrew characters, because of common educational and literacy practices. Orthographic practices have varied, especially in the representation of vowels.

Jewish languages are often strongly influenced by a language spoken by the group's ancestors. In the case of Yiddish, the main previous Jewish language was Judeo-French. In the case of Judeo-Spanish, the main previous Jewish language was Judeo-Arabic. And in the case of Jewish English, the main previous Jewish language was Yiddish. These previous languages provide influences in lexicon, as well as other areas. In addition, the previous languages have a major impact on the use of Hebrew and Aramaic: which words are used, how they are pronounced, and how they are integrated morpho-syntactically.

Most Jewish communities have used the local language in distinctive ways in their translations of biblical and liturgical texts. These translations tend to render the local lexicon in word-for-word imitations of the Hebrew syntax. This practice is referred to in various ways, e.g., Judeo-Arabic *Sharḥ*, Yiddish *Taytsh*, and Judeo-Spanish *Ladino*.

The revitalization and re-vernacularization of Hebrew as part of the Zionist enterprise have produced a new situation, where modern Israeli Hebrew is markedly distinct from its earlier forms. Is this new variety to be considered a 'Jewish language'? Some have argued that it is too different from Diaspora Jewish languages to be classified with them. On the other hand, it shares many features: strong influence of the previous Jewish language (the Yiddish base is most evident in the highly modified grammar), a special place for Hebrew-Aramaic lexical items, borrowing from the co-territorial non-Jewish languages (including spoken Arabic and the widely known English), and Hebrew orthography.

The Most Widely Spoken Jewish Language: Yiddish

Distinctive linguistic features can be seen in the history of Yiddish, the most widely spoken Jewish language (Birnbaum, 1979; Weinreich, 1980; Katz, 1985; Weigel, 2002). According to the commonly accepted view (Weinreich, 1980), Yiddish was born towards the end of the first millennium A.D. when Judeo-French-speaking Jews started to settle in the Rhineland. During the more tolerant period that preceded the Crusades, these communities shifted from Judeo-French to a variety based on the German spoken in the area. This Judeo-German included elements of Hebrew and Aramaic, as well as other distinctive features. As a result of expulsion, persecution, and changing economic opportunity, many Jews migrated from Germanic-speaking to Slavic-speaking areas and brought their German language with them. In the changed social conditions, the developing Yiddish language maintained its German base while admitting influences from local Slavic languages in lexicon, morphosyntax, phonology, and discourse. In addition, Hebrew and Aramaic elements survived from previous generations, and new ones were added, mostly through contact with liturgical and rabbinic texts. A few Judeo-French lexical elements endured.

Over the centuries, Western Yiddish disappeared as a spoken language, assimilating towards co-territorial German, except in a few areas, like French-speaking Alsace. In eastern Europe, Yiddish developed into a complex web of dialects, differing mostly in phonology but also in lexicon and grammar.

Yiddish documents have been identified as early as the 13th century, and we have examples of epic poems written in Yiddish from the time of the Renaissance. In the early modern period, Yiddish was used mostly in women's religious literature, including translations and explanations of the Bible and liturgy. The mid-19th century saw the flowering of Yiddish literature, stemming from the eastern European Jewish Enlightenment. In the early 20th century, Yiddish became the object of language planning efforts, including a standardized orthography and linguistic documentation and research.

Events of the 20th century, especially immigration to America and Israel and the Nazi Holocaust, led to a major decline in the use of Yiddish, and today it is used as an everyday language mostly by the elderly and by pockets of Hasidic Jews in the New York area, Israel, and elsewhere (current estimates of total number of speakers range from 200 000 to 400 000). Reversing language shift efforts continue in educational and cultural programs, especially in New York, Montreal, Antwerp, and Mexico City. Young non-Hasidic Jews there and elsewhere continue to use Yiddish as an everyday language in an effort at revitalization.

Judeo-Spanish

Judeo-Spanish is an Hispanic language taken by exiles from Spain after the expulsion of 1492 to northern Europe, the Balkans and Turkey (Sephiha, 1979; Malinowski, 1982; Bunis, 1993; Harris, 1994; Quintana, 2002). There has been much debate about the name of this language: in addition to Judeo-Spanish, commonly used glottonyms are Ladino, Judezmo, and Spanyol, with some scholars maintaining that Ladino should refer only to the calque (word-for-word) translation language variety.

Already in Spain, Judeo-Spanish exhibited influences from Jewish and non-Jewish varieties of Arabic, as well as other distinctive features. When Sephardic Jews migrated, elements of Turkish, Greek, Bulgarian, and other languages were added. In addition, archaisms and independent developments distinguished

Judeo-Spanish from contemporaneous peninsular Castilian. Distinctive dialects of Judeo-Spanish formed throughout the Ottoman Empire. In the 19th and 20th centuries, the high-status French language had a major impact, due to influences of religious and secular education.

Judeo-Spanish developed literary functions, including a significant religious literature, a strong oral folk literature, and a corpus of modern belles lettres. There was rapid language loss in the 19th century as a result of emigration, Westernization, and assimilation. The Judeo-Spanish-speaking community in Greece and other Balkan countries was mostly wiped out in the Holocaust. Speakers in Turkey shifted first to French and more recently to Turkish. Today, it is estimated that there are only 30 000–50 000 speakers, mostly elderly.

A North African variety of Judeo-Spanish, called Haketiya, developed in northern Morocco after the 1492 expulsion. Its speakers mostly shifted to Spanish with the establishment of the Spanish Protectorate at the beginning of the 20th century.

Judeo-Arabic

Since even before the Muslim conquest of the Arabian peninsula in the 7th century A.D., Jews have lived alongside Arabic speakers, and they have spoken Jewish varieties of Arabic (Blanc, 1964; Blau, 1981; Hary, 1992; Bar-Asher, 1998). These have included Hebrew and Aramaic influences – mostly lexical, but also phonological, morphological, and syntactic. They have also included archaisms, standardized hyper- and hypo-corrections, and other distinctive features. Judeo-Arabic varieties have been documented in Iraq, Egypt, Syria, Morocco, and Yemen. Due to the migration of Jews within the Arab world, some varieties have features in common with other varieties of Judeo-Arabic that do not exist in the local non-Jewish Arabic dialects.

In the Middle Ages, many important Jewish religious and philosophical works were written in Middle Arabic and Judeo-Arabic. A word-for-word, or calque, translation variety, called *Sharḥ*, was used for translations of biblical, rabbinic, and liturgical texts. In addition, Judeo-Arabic was used for religious and secular literary production in the 19th century.

Most Jews in Arab lands immigrated to Israel in the 20th century, acquiring Israeli Hebrew and relegating Judeo-Arabic to private use. Those who stayed in Morocco tended to shift to French, and those who immigrated to North America and France tended to shift to the local languages there. It is estimated that there are currently 400 000–500 000 speakers (Grimes, 1996), mostly middle aged and older.

A Contemporary Jewish Language: Jewish English

Also referred to as Judeo-English, Yinglish, and Yeshivish, Jewish English is an umbrella term for the contemporary in-group varieties spoken by Jews in America, England, and other English-speaking countries (Gold, 1985; Steinmetz, 1986; Weiser, 1994; Benor, 2004). Jewish English is based on the local variety of English with many influences from Yiddish, textual Hebrew and Aramaic, and Israeli Hebrew in lexicon, syntax, phonology, and discourse. Because of the widespread literacy in contemporary English-speaking countries, Jewish English is not written in Hebrew characters. However, Hebrew loan words are sometimes inserted in their original orthography.

The varieties of Jewish English spoken by Orthodox Jews, especially those in larger, more isolated communities, are most distinct from general English, often to the point of being unintelligible to non-Jews and to non-Orthodox Jews. Orthodox Jewish English is a young language variety, and the number of speakers is growing as the Orthodox community expands. Although Jewish English is the youngest Jewish language that has been researched, it is likely not the only one that is gaining, rather than losing, speakers. As researchers explore the language of other contemporary Diaspora communities, it is expected that they will find similar distinctively Jewish linguistic practices.

See also: Bilingualism; Code Switching; Hebrew, Biblical and Jewish; Hebrew, Israeli; Language Change and Language Contact; Language Maintenance and Shift; Minorities and Language; Yiddish.

Bibliography

Benor S B (2004). 'Second style acquisition: the linguistic socialization of newly Orthodox Jews.' Ph.D. diss., Stanford University.

Bar-Asher M (1998). *Linguistic traditions of North African Jews.* Jerusalem: Hebrew University. [In Hebrew].

Birnbaum S A (1979). *Yiddish: a survey and a grammar.* Toronto: University of Toronto Press.

Blanc H (1964). *Communal dialects in Baghdad.* Cambridge, MA: Harvard University Press.

Blau J (1965). *The emergence and linguistic background of Judeo-Arabic: a study of the origins of Middle Arabic.* London: Oxford University Press.

Blau J (1981). *The emergence and linguistic background of Judaeo-Arabic.* Jerusalem: Ben-Zvi Institute.

Bogoch B (1999). 'Gender, literacy and religiosity: dimensions of Yiddish education in Israeli government-supported schools.' *International Journal of the Sociology of Language 138,* 123–160.
Bunis D M (1993). *A lexicon of the Hebrew and Aramaic elements in modern Judezmo.* Jerusalem: Magnes Press & Misgav Yerushalayim.
Chomsky W (1957). *Hebrew: the eternal language.* Philadelphia: The Jewish Publication Society of America.
Fellman J (1973). *The revival of a classical tongue: Eliezer ben Yehuda and the modern Hebrew language.* The Hague: Mouton.
Fishman J A (1980). 'Attracting a following to high-culture functions for a language of everyday life: the role of the Tshernovits (Conference in the 'Rise of Yiddish').' *International Journal of the Sociology of Language 24,* 43–73.
Fishman J A (ed.) (1985). *Readings in the sociology of Jewish languages.* Leiden: E. J. Brill.
Garbell I (1965). *The Jewish Neo-Aramaic dialect of Persian Azerbaijan.* The Hague: Mouton.
Gold D (1985). 'Jewish English.' In Fishman J A (ed.) *Readings in the sociology of Jewish languages.* Leiden: Brill. 280–298.
Gold D (1989). *Jewish linguistic studies.* Haifa: Association for the Study of Jewish Languages.
Gold D & Prager L (eds.) (1981–1987). *Jewish language review.* Haifa: Association for the Study of Jewish Languages.
Greenfield J C (1978). 'Aramaic and its dialects.' In Paper H H (ed.) *Jewish languages: themes and variations.* Cambridge, MA: Association for Jewish Studies. 29–43.
Grimes B A (ed.) (1996). *Ethnologue: languages of the world* (13th edn.). Dallas, Texas: Summer Institute of Linguistics.
Harris T K (1994). *Death of a language: the history of Judeo-Spanish.* Newark: University of Delaware Press.
Harshav B (1993). *Language in time of revolution.* Berkeley: University of California Press.
Hary B (1992). *Multiglossia in Judeo-Arabic: with an edition, translation and grammatical study of the Cairene Purim scroll.* Leiden: Brill.
Isaacs M (1999). 'Contentious partners: Yiddish and Hebrew in Haredi Israel.' *International Journal of the Sociology of Language 138,* 101–121.
Jewish Language Research Website: http://www.jewish-languages.org/.
Katz D (1985). 'Hebrew, Aramaic and the rise of Yiddish.' In Fishman J A (ed.) *Readings in the sociology of Jewish languages.* Leiden: E. J. Brill. 85–103.
Malinowski A (1982). 'A report on the status of Judeo-Spanish in Turkey.' *International Journal of the Sociology of Language 37,* 7–23.
Mieses M (1915). *Die Entstehungsursache der jüdischen Dialekte.* Vienna: R. Löwit Verlag.
Paper H H (ed.) (1978). *Jewish languages: themes and variations.* Cambridge, MA: Association for Jewish Studies.
Pilowsky A L (1985). 'Yiddish alongside the revival of Hebrew: public polemics on the status of Yiddish in Eretz Israel, 1907–1929.' In Fishman J A (ed.) *Readings in the sociology of Jewish Languages.* Leiden: Brill. 104–124.
Quintana A (2002). 'Geografía lingüística del Judeo-español de acuerdo con el léxico.' *Revista de Filología Española 82(1–2),* 105–138.
Rabin C (1981). 'What constitutes a Jewish language?' *International Journal of the Sociology of Language 30,* 19–28.
Rabin C *et al.* (1979). 'The Jewish languages: commonalities, differences and problems (in Hebrew).' *Pe'amim 1,* 40–57.
Sabar Y (2002). *A Jewish Neo-Aramaic dictionary: dialects of Amidya, Dihok, Nerwa and Zakho, northwestern Iraq.* Wiesbaden: Harrassowitz.
Sephiha H V (1979). *Le Ladino (Judeo-Espagnol calque).* Paris: Association Vidas Largas.
Spolsky B (1991). 'Hebrew language revitalization within a general theory of second language learning.' In Cooper R L & Spolsky B (eds.) *The influence of language on culture and thought: essays in honor of Joshua A. Fishman's sixty-fifth birthday.* Berlin: Mouton de Gruyter. 137–155.
Spolsky B & Shohamy E (1999). *The languages of Israel: policy, ideology and practice.* Clevedon: Multilingual Matters.
Steinmetz S (1987). *Yiddish and English: a century of Yiddish in America.* University: University of Alabama Press.
Weigel W (2002). 'Yiddish.' *Jewish language research website.* http://www.jewish-languages.org/yiddish.html.
Weinreich M (1980). *History of the Yiddish language.* Noble S & Fishman J (trans.). Chicago: University of Chicago Press.
Weiser C (1995). *Frumspeak: the first dictionary of Yeshivish.* Northvale, NJ: Jason Aronson.
Wexler P (1981). 'Jewish interlinguistics: facts and conceptual framework.' *Language 57,* 99–149.
Wexler P (1985). 'Recovering the dialects and sociology of Judeo-Greek in non-Hellenic Europe.' In Fishman J A (ed.) *Readings in the sociology of Jewish languages.* Leiden: E. J. Brill. 227–240.

Relevant Website

http://www.jewish-languages.org – The Jewish Language Research Website.

Jiwarli

P K Austin, School of Oriental and African Studies, London, UK

Introduction

Jiwarli (Djiwarli) is an Australian Aboriginal language and was traditionally spoken along the upper reaches of the Henry River, a tributary of the Ashburton River, in the northwest of Western Australia. The language was unrecorded until 1978 and is now extinct, the last speaker, Mr. Jack Butler, having passed away in May 1985. Before his death, Jack Butler worked with Peter Austin to record over 70 texts in a range of genres, a lexicon of some 1500 words and elicitation of morphological paradigms and syntactic constructions. Publications on the language include a bilingual dictionary (Austin, 1992), a text collection (Austin, 1997), articles on morphosyntax (Austin and Bresnan, 1996; Austin, 1995, 1998, 2000, 2001) and a website. The language has become known in the linguistic literature for its nonconfigurational syntax (see Austin and Bresnan, 1996; Baker, 2000, and below), and it also shows switch-reference and a complex system of case-marking that reflects clause type (Austin, 2004). A reference grammar is in preparation.

Language Relationships

Jiwarli is closely related to its immediate neighbors, Warriyangka (Warriyangga), Thiin, and Tharrkari (Dhargari) as members of the Mantharta subgroup (*mantharta* being the word for 'person'). The languages share up to 80% common vocabulary and a similar grammatical system. Tharrkari has undergone a number of historical phonological changes that make its phonetics and phonology highly unusual for an Australian language (see further below). None of the Mantharta languages has any native speakers today, though some knowledge of words and expressions remains among descendants. The Mantharta languages are most closely related to the Kanyara languages spoken to their west and northwest: Payungu (Bayungu), Pinikura (Binigura), Purduna (Burduna), and Thalanyji (Dhalandji). They share approximately 60% cognate vocabulary and have a number of grammatical features in common, including switch-reference and clause linkage effects on case-marking (Austin, 1996, 2004). Today only Thalanyji continues to be spoken by older members of a single family living near Onslow, Western Australia. The Kanyara and Mantharta languages belong to the widespread Pama–Nyungan family, which covers the southern two-thirds of Australia (*see* **Australia: Language Situation**), and are most closely related to the Nyungic languages spoken to their east.

Linguistic Characteristics

Phonology

The phonological system of Jiwarli is typical of languages of the region, with contrastive stops at six points of articulation, a nasal for each stop position, a lateral for each nonperipheral stop, a flap, a semi-retroflex continuant, and two glides. Table 1 gives the relevant consonants in their practical orthographic form. There are just three vowels: high front *i*, high back *u*, and low *a*, with a phonemic length contrast mainly, but not exclusively, found in the first syllable of words. Tharrkari has undergone a number of historical phonological changes that have resulted in the creation of a stop voicing contrast (unusual in Australia) and, in one dialect, complete loss of laterals.

The general structure of Jiwarli roots is CV(C)CV(C). Every word in Jiwarli must begin with a consonant and end in a vowel; roots can end in a consonant but if otherwise unsuffixed *-ma* is added to nasal-final roots and *-pa* to roots ending in *l*, *rl*, *ly*, or *rr*. Word-initially only nonapico-domal stops and nasals and the two glides are found. Word-medially

Table 1 Consonants

	Bilabial	*Lamino-*		*Apico-*		*Dorse-velar*
		dental	*palatal*	*alveolar*	*domal*	
Stop	*p*	*th*	*j*	*t*	*rt*	*k*
Nasal	*m*	*nh*	*ny*	*n*	*rn*	*ng*
Lateral		*lh*	*ly*	*l*	*rl*	
Flap				*rr*		
Continuant					*r*	
Glide	*w*		*y*			

Table 2 Substantive cases

	A	*S*	*P*	*Dative*	*Locative*	*Allative*
boy	*wirtangku*	*wirta*	*wirtanha*	*wirtawu*	*wirtangka*	*wirtarla*
girl	*kurlkingku*	*kurlki*	*kurlkinha*	*kirlkiyi*	*kurlkingka*	*kurlkirla*
dog	*thuthungku*	*thuthu*	*thuthunha*	*thuthuwu*	*thuthungka*	*thuthurla*
fire	*karlangku*	*karla*	*karla*	*karlawu*	*karlangka*	*karlarla*
tree	*wurungku*	*wuru*	*wuru*	*wuruwu*	*wurungka*	*wururla*
hill 'roo	*mathantu*	*mathanma*	*mathannha*	*mathanku*	*mathanta*	*mathankurla*
tongue	*thalanythu*	*thalanyma*	*thalanyma*	*thalanyku*	*thalanytha*	*thalanykurla*
chin	*nyinyarntu*	*nyinyarnma*	*nyinyarnma*	*nyinyarnku*	*nyinyarnta*	*nyinyarnkurla*
wind	*yuwalpalu*	*yuwalpa*	*yuwalpa*	*yuwalku*	*yuwalpala*	*yuwalkurla*
cousin	*ngathalpalu*	*ngathalpa*	*ngathalpanha*	*ngathalku*	*ngathalpala*	*ngathalkurla*
barb	*ngarlirrpalu*	*ngarlirrpa*	*ngarlirrpa*	*ngarlirrku*	*ngarlirrpalu*	*ngarlirrkurla*

there are limited consonant clusters, primarily homorganic nasal plus stop, and apical nasal or lateral plus peripheral stop (*p* and *k*). Vowel clusters are not found. Words borrowed from English are generally restructured to meet these phonotactic constraints, e.g., *walypala* 'white man' (from 'white fella'), *ngayirlanma* 'island.'

Morphology

Jiwarli, like other languages of the Pama–Nyungan group, is entirely suffixing in its morphology. There are two major word classes, nominals and verbs, with nominals showing a rich system of case-marking and verbs marking tense/aspect/mood and dependent clause categories. Nominals can be subdivided into substantives (which cover both noun and adjective concepts in a language like English), pronouns, locationals, and demonstratives. Minor word classes include adverbs, particles, and interjections.

Nominals in Jiwarli inflect for case, with the syntactic functions of intransitive subject (S), transitive subject (A), and transitive object (P) showing a split-ergative pattern of syncretism in the case forms determined by animacy:

- for the first person singular pronoun, S and A fall together as a single (unmarked) form
- for inanimate nominals and demonstratives, S and P fall together as a single (unmarked) form
- for all other nominals, there are three forms for S, A, and P functions

In addition to the three main cases (nominative for S, ergative for A, accusative for P) there are also the following case forms:

- dative, marking alienable possession and complement of certain verbs
- allative, coding direction toward a place
- locative, coding location in a place, and complement of verbs of speaking
- ablative, coding direction from a place and cause

The actual forms of the cases are affected by the phonological shape of the root, e.g., whether it ends in a vowel or not, what kind of vowel or consonant is root-final, and how many morae it contains (long vowels counting as two mora). Table 2 sets out a sample substantive declension.

The coding of transitive object P varies according to clause type and crossclausal reference relations: in certain dependent clauses (for details, see below) P is marked as dative or as allative. In addition, case is added to dependent clause verbs to indicate cross-clausal coreference (see below), and manner adverbs in transitive clauses carry an ergative case agreement marker.

Jiwarli has a rich system of nominal word-building morphology that involves suffixation between the root and case inflection. Categories encoded in word-building morphology include number (dual, paucal, plural), having (e.g., *yakan-jaka* 'married [lit. spouse-having]'), lacking (e.g., *yakan-yirra* 'unmarried [lit. spouse-lacking]'), and kin dual and plural (e.g., *kurta* 'older brother,' *kurtarra* 'pair of brothers').

Pronouns in Jiwarli distinguish three persons and singular, dual, and plural number; in the first person non-singular there is an inclusive–exclusive contrast. Demonstratives encode a proximal and distal contrast. Table 3 sets out the basic pronoun and demonstrative forms.

Verbs morphologically distinguish between main verb and dependent verb inflections. Main verbs encode tense/aspect/mood categories such as past habitual, present, future, and imperative. Dependent verbs occur in hypotactically linked clauses and encode clause type (relative tense plus aspect) plus cross-clausal coreference or non-coreference of subjects (S or A), i.e., switch-reference (see further details in the syntax section below). There are five morphologically determined verb conjugations: conjugations one and two are primarily, but not exclusively, transitive,

Table 3 Pronouns and demonstratives

	A	*S*	*P*	*Dative*	*Locative*
1sg	*ngatha*	*ngatha*	*ngathanha*	*nganaju*	*ngathala*
1dlincl	*ngalilu*	*ngali*	*ngalinha*	*ngalimpa*	*ngalila*
1dlexcl	*ngalijuru*	*ngaliju*	*ngalijunha*	*ngalijungu*	*ngalijura*
1plincl	*nganthurralu*	*nganthurru*	*nganthurranha*	*nganthurrampa*	*nganthurrala*
1plexcl	*nganthurrajuru*	*nganthurraju*	*nganthurrajunha*	*nganthurrajungu*	*nganthurrajura*
2sg	*nhurralu*	*nhurra*	*nhurranha*	*nhurrampa*	*nhurrala*
2dl	*nhupaluru*	*nhupalu*	*nhupalunha*	*nhupalumpa*	*nhupalura*
2pl	*nhurrakaralu*	*nhurrakara*	*nhurrakaranha*	*nhurrakarampa*	*nhurrakarala*
3sg	*panhaluru*	*panhalu*	*panhalunha*	*parnumpa*	*panhalura*
3dl	*pulalu*	*pula*	*pulanha*	*pulampa*	*pulala*
3pl	*thanalu*	*thana*	*thananha*	*thanampa*	*thanala*
this	*yilu*	*yinha*	*yinha*	*yirnu*	*yila*
that	*ngulu*	*ngunha*	*ngunha*	*ngurnu*	*ngula*

Table 4 Verb inflections

Inflection	*Conj 1*	*Conj 2*	*Conj 3*	*Conj 4*	*Conj 5*
Main clause verb inflections					
Usitative	*-laartu*	*-rraartu*	*-artu*	*-artu*	*-artu*
Past	*-rninyja*	*-rninyja*	*-nyja*	*-nyja*	*-nyja*
Present	*-nha*	*-nha*	*-inha*	*-inha*	*-a*
Future	*-lka*	*-rrka*	*-ira*	*-ra*	*-ra*
Imperative	*-nma*	*-nma*	*-ma*	*-ma*	*-ma*
Irrealis	*-nmararni*	*-nmararni*	*-mararni*	*-mararni*	*-mararni*
Dependent clause verb inflections					
ImperfSS	*-rnu*	*-rnu*	*-nhu*	*-nhu*	*-nhu*
ImperfDS	*-niya*	*-niya*	*-iniya*	*-iniya*	*-iniya*
PerfSS	*-rninyjalu*	*-rninyjalu*	*-nyjalu*	*-nyjalu*	*-nyjalu*
PerfDS	*-rninyjaparnti*	*-rninyjaparnti*	*-nyjaparnti*	*-nyjaparnti*	*-nyjaparnti*
PurpSS	*-ru*	*-rru*	*-yi*	*-ngku*	*-a*
PurpDS	*-lpuka*	*-rrpuka*	*-puka*	*-puka*	*-puka*
Intentive	*-lkarri(ngu)*	*-rrkarri(ngu)*	*-irarri(ngu)*	*-rarri(ngu)*	*-rarri(ngu)*
Might	*-lkangu*	*-rrkangu*	*-irangu*	*-rangu*	*-rangu*

and conjugations three, four, and five are intransitive. Table 4 sets out the verb conjugations.

Verbs show limited word-building morphology, mainly transitivizing and detransitivizing affixes that shift conjugation and transitivity. There are also category-changing affixes:

- nominalizing suffixes that create agent or instrument nominals from verbs
- verbalizing suffixes that create intransitive (inchoative) or transitive (causative) verbs from nominals.

The minor categories of adverb, particle, and interjection show no morphological variation. However, there is a set of postinflectional suffixes that may be attached to words of any class to encode various information status concepts, such as *-rru* for 'new information' and *-thu* for 'old information.' These affixes are ubiquitous in texts.

Syntax

Jiwarli is a nonconfigurational language (Hale, 1983; Austin and Bresnan, 1996; Baker, 2000) and shows the following syntactic characteristics:

- free word order, in which any possible order of sentence constituents is found (Austin, 2000)
- split-NP syntax, in which nominals understood as referring to a single entity can be separated in the clause by other constituents (each nominal bearing a relevant case marker)
- free argument ellipsis, in which nominals of any person or number whose reference is clear from the context can be freely omitted (Austin, 2001).

The following example illustrates split-NP syntax:

(1) *Karla* *wantha-nma-rni* *jarnpa* *juma*
fire.ACC give-IMPER-hence light.ACC small.ACC
'Give me a small fire light' [T52s15]

Free elipsis of arguments is seen in the following (see also line 68 in the text example below):

(2) *Wirntupinya-nyja-rru*
kill-PAST-NEW.INF
'(They) killed (him)' [T42s25]

Jiwarli also shows interesting interclausal syntax. Dependent clauses occur hypotactically located on the margins of main clauses, and their verbs encode clause type plus switch-reference, i.e., (non-)coreference of subject (S or A) between the main and dependent clause. In same-subject clauses the dependent clause subject is obligatorily unexpressed (these being 'control' structures). When the main clause is transitive, some same-subject dependent clause verbs carry an ergative case marker in agreement with the controlling subject nominal. The following examples illustrate this:

(3) *Mantharta* *kumpa-inha* *wurnta-wu*
man.NOM sit-PRES shield-DAT
yinka-rnu
adze-IMPERF.SS
'The man sits adzing a shield' [N11p31s3]

(4) *Nhurra-kara-lu* *thika-nma*
you-PL-ERG eat-IMPER
yarrukarri-ngu-ru-thu
want-IMPERF.SS-ERG-old.INF
'You eat it if you want it!' [N11p39s3]

For different-subject dependent clauses, if there is coreference between the (omitted) subject of the dependent clauses and a nonsubject in the main clause, an agreement case marker appears on the dependent verb, as in:

(5) *Tharla-nma* *yinha* *julyu-nha*
feed-IMPER this.ACC old man-ACC
kamu-rri-ya-nha
hunger-INCHOAT-IMPERF.DS-ACC
'Feed this old man who is becoming hungry!' [JIT13s1]

Notice that there is a complex interaction between the marking of P inside the dependent clause (as dative, allative, or accusative) depending on clause type and crossclausal coreference relations. Table 5 illustrates this.

Table 5 Coding of P

	Dependent object
Intentive	Dative
Imperfective-SS	Dative
Perfective-SS	Dative
Imperfective-DS	Dative
Perfective-DS	Dative
Purposive-SS	Allative
Purposive-DS	Accusative
Might	Accusative

The significance of these patterns is explored more generally in Austin (2004).

Particles in Jiwarli have scope over the whole clause and encode such semantic concepts as negation, possibility, etc. An example is *warri* 'not' in (see also *kaji* 'try' in text line 15 below):

(6) *Nhaa-rri-nyja* *nhurra* *warri*
what-INCHOAT-PAST you.ERG not
kurlkayi-rnu *wangka-iniya-wu* *nganaju*
listen-IMPERF.SS talk-IMPERF.DS-DAT I.DAT
'Why didn't you listen to me talking?' [T35s7]

Text Example

The following extract (Text 43 in Austin, 1997) exemplifies the morphological and syntactic characteristics of Jiwarli and shows a little of the cultural background of the language. It comes from a traditional story in which one bird steals fire from the people, who then ask the Peregrine falcon to get it back from the thief:

(7) *Ngana-lu* *ngunha* *karla* *mana-ra*
who-ERG that.ACC fire.ACC get-FUT
'Who will get the fire?'

(8) *Nhurra* *parru*
2SG.NOM then
'How about you?'

(9) *Nhurra* *karlathintirnira* *kurukurura,*
2SG.NOM Peregrine falcon Peregrine falcon
nhurra *yini-thu*
2SG.NOM name.NOM-OLD.INF
'You are karlathintirnira Peregrine falcon, (that's) your name.'

(10) *Ngaa*
yes
'Yes'

(11) *Ngunha* *thurni-nyja-nthi*
that.NOM laugh-PAST-just
'He just laughed'

(12) *Yana-nyja* *ngunha* *purtipala-rru*
go-PAST that.NOM pretty.NOM-NEW.INF
'He was pretty now'

(13) *Wantha-rninyja* *juuri* *wangkarr-a*
put-PAST paint.ACC throat-LOC
'(They) put paint on (his) throat'

(14) *Wantha-rninyja* *kala-pa* *wangkarr-a*
put-PAST like this-SPEC throat-LOC
'(They) put (it) like this on (his) throat'

(15) *Kaji* *nhurra* *yana-ma* *mana-ngku*
try 2SG.NOM go-IMPER get-PURP.SS
ngurlu *karla-rla*
that.ALL fire-ALL
'You try to go and get the fire'

See also: Australia: Language Situation.

Bibliography

Austin P (1992). *A dictionary of Jiwarli, Western Australia.* Melbourne: La Trobe University.

Austin P (1995). 'Double case marking in Kanyara and Mantharta languages.' In Plank F (ed.) *Agreement by Suffixaufnahme*. Oxford: Oxford University Press. 363–379.

Austin P (1996). 'Kanyara and Mantharta languages.' University of Melbourne. [Prepared for Australian Linguistic Institute Workshop, 'Where did the Western Desert language come from?' convened by Patrick McConvell (NTU) and Mary Laughren (UQld).]

Austin P (1997). *Texts in the Mantharta languages, Western Australia.* Tokyo: Tokyo University of Foreign Studies.

Austin P (1998). 'Eaglehawk was sitting chasing them: grammaticisation of the verb 'to sit' in Mantharta languages, Western Australia.' In Siewierska A & Jung Song J (eds.) *Case, typology and grammar: in honour of Barry J. Blake*. Typological Studies in Language 38. Amsterdam: John Benjamins.

Austin P (2000). 'Word order in a free word order language: the case of Jiwarli.' In Simpson J, Nash D, Laughren M, Austin P & Alpher B (eds.) *Forty years on: Ken Hale and Australian languages.* Canberra: Pacific Linguistics. 305–324.

Austin P (2001). 'Zero arguments in Jiwarli, Western Australia.' *Australian Journal of Linguistics 21(1),* 83–98.

Austin P (2004). 'Case and clauses in Australian Aboriginal languages.' University of London, manuscript.

Austin P & Bresnan J (1996). 'Non-configurationality in Australian Aboriginal languages.' *Natural Language and Linguistic Theory 14,* 215–268.

Baker M (2001). 'Non-configurationality.' In Baltin M & Collins C (eds.) *The handbook of syntactic theory.* London: Blackwells.

Hale K (1983). 'Warlpiri and the grammar of non-configurational languages.' *Natural Language and Linguistic Theory 1(1),* 5–74.

Johnson, Samuel (1709–1784)

W Hüllen, Düsseldorf, Germany

Samuel Johnson was born as the son of a bookseller in Lichfield in 1709 and died in London in 1784. He was an intellectually precocious boy, devouring literature in a desultory manner, and later a young man of often unpredictable habits and uncouth appearance. This was to become his lifestyle, which alternated with lapses into deep melancholy. In spite of the irregularities of his character, his deep sympathy for underprivileged persons and servants was unanimously praised. At Oxford (Pembroke College), he read Greek and 'metaphysics.' In 1729, he left the university before obtaining a degree, probably because his means failed. (The degree of LL.D., which later became habitually attached to his name, was conferred upon him by the University of Dublin as late as 1765 and by the University of Oxford even later in 1775. He rarely used it himself.) After several attempts at gaining a position at schools in the neighborhood of Lichfield, he left for London in 1737, where he stayed for the rest of his life.

Samuel Johnson was determined to make a living as an *homme de lettres* and, after enduring utmost poverty for many years, finally succeeded in doing so. His literary works were of only limited success. Today they are regarded as fine, though obsolete, specimens of classicist literature. Apart from their formal achievements, they gave rise to the general appreciation of Samuel Johnson as a morally strong and religiously motivated character. Among them were the tragedy *Irene* (1736, staged by Garrick in 1749), the long poem 'Vanity of Human Wishes,' written in the imitation of the tenth satire of Juvenal (1749), and the novel *Rasselas, Prince of Abyssinia* (1759). All of them quote historical examples of the deceptive nature of human achievements and the everlasting and unique salvation in God. They propose this message using a severe and dignified eloquence that was felt to be exemplary.

More successful were Johnson's *ad hoc* publications (prologues to plays, reviews, political pamphlets) and his contributions to various journals. From 1731 on he wrote regularly for the *Gentlemen's Magazine*. Among his contributions written between 1738 and 1744 were parliamentary debates, some of which he only corrected or wrote on the basis of reports; he was never in the gallery himself. Between 1750 and 1752, he issued the *Rambler*, which appeared twice a week, by writing the whole paper himself with a few exceptions. The project was not successful but later gained a high reputation and good sales when it was edited in collected volumes. In 1753 and 1754, he contributed to the *Adventurer*. Between 1756 and 1758, he edited the *Literary Magazine, or Universal Review*, again writing a good many of the articles

See also: Australia: Language Situation.

Bibliography

Austin P (1992). *A dictionary of Jiwarli, Western Australia.* Melbourne: La Trobe University.

Austin P (1995). 'Double case marking in Kanyara and Mantharta languages.' In Plank F (ed.) *Agreement by Suffixaufnahme.* Oxford: Oxford University Press. 363–379.

Austin P (1996). 'Kanyara and Mantharta languages.' University of Melbourne. [Prepared for Australian Linguistic Institute Workshop, 'Where did the Western Desert language come from?' convened by Patrick McConvell (NTU) and Mary Laughren (UQld).]

Austin P (1997). *Texts in the Mantharta languages, Western Australia.* Tokyo: Tokyo University of Foreign Studies.

Austin P (1998). 'Eaglehawk was sitting chasing them: grammaticisation of the verb 'to sit' in Mantharta languages, Western Australia.' In Siewierska A & Jung Song J (eds.) *Case, typology and grammar: in honour of Barry J. Blake.* Typological Studies in Language 38. Amsterdam: John Benjamins.

Austin P (2000). 'Word order in a free word order language: the case of Jiwarli.' In Simpson J, Nash D, Laughren M, Austin P & Alpher B (eds.) *Forty years on: Ken Hale and Australian languages.* Canberra: Pacific Linguistics. 305–324.

Austin P (2001). 'Zero arguments in Jiwarli, Western Australia.' *Australian Journal of Linguistics 21(1),* 83–98.

Austin P (2004). 'Case and clauses in Australian Aboriginal languages.' University of London, manuscript.

Austin P & Bresnan J (1996). 'Non-configurationality in Australian Aboriginal languages.' *Natural Language and Linguistic Theory 14,* 215–268.

Baker M (2001). 'Non-configurationality.' In Baltin M & Collins C (eds.) *The handbook of syntactic theory.* London: Blackwells.

Hale K (1983). 'Warlpiri and the grammar of non-configurational languages.' *Natural Language and Linguistic Theory 1(1),* 5–74.

Johnson, Samuel (1709–1784)

W Hüllen, Düsseldorf, Germany

Samuel Johnson was born as the son of a bookseller in Lichfield in 1709 and died in London in 1784. He was an intellectually precocious boy, devouring literature in a desultory manner, and later a young man of often unpredictable habits and uncouth appearance. This was to become his lifestyle, which alternated with lapses into deep melancholy. In spite of the irregularities of his character, his deep sympathy for underprivileged persons and servants was unanimously praised. At Oxford (Pembroke College), he read Greek and 'metaphysics.' In 1729, he left the university before obtaining a degree, probably because his means failed. (The degree of LL.D., which later became habitually attached to his name, was conferred upon him by the University of Dublin as late as 1765 and by the University of Oxford even later in 1775. He rarely used it himself.) After several attempts at gaining a position at schools in the neighborhood of Lichfield, he left for London in 1737, where he stayed for the rest of his life.

Samuel Johnson was determined to make a living as an *homme de lettres* and, after enduring utmost poverty for many years, finally succeeded in doing so. His literary works were of only limited success. Today they are regarded as fine, though obsolete, specimens of classicist literature. Apart from their formal achievements, they gave rise to the general appreciation of Samuel Johnson as a morally strong and religiously motivated character. Among them were the tragedy *Irene* (1736, staged by Garrick in 1749), the long poem 'Vanity of Human Wishes,' written in the imitation of the tenth satire of Juvenal (1749), and the novel *Rasselas, Prince of Abyssinia* (1759). All of them quote historical examples of the deceptive nature of human achievements and the everlasting and unique salvation in God. They propose this message using a severe and dignified eloquence that was felt to be exemplary.

More successful were Johnson's *ad hoc* publications (prologues to plays, reviews, political pamphlets) and his contributions to various journals. From 1731 on he wrote regularly for the *Gentlemen's Magazine*. Among his contributions written between 1738 and 1744 were parliamentary debates, some of which he only corrected or wrote on the basis of reports; he was never in the gallery himself. Between 1750 and 1752, he issued the *Rambler*, which appeared twice a week, by writing the whole paper himself with a few exceptions. The project was not successful but later gained a high reputation and good sales when it was edited in collected volumes. In 1753 and 1754, he contributed to the *Adventurer*. Between 1756 and 1758, he edited the *Literary Magazine, or Universal Review*, again writing a good many of the articles

himself. All this meant accomplishing an enormous amount of work, even during the period when he was busy compiling his dictionary. The pecuniary conditions of his life improved considerably when he was granted a pension by George III in 1762.

Samuel Johnson then became an outstanding figure of the intellectual scene in London. Nearly every distinguished man of letters of the period came into contact with him, with the exception of the disliked David Hume (Dictionary of National Biography, DNB). In 1763, he made the acquaintance of James Boswell, and started a friendship that lasted even beyond the grave, as his friend published the well-known and exemplary biography *Life of Samuel Johnson* (1791). In the same year, the 'Club' was founded – a regular gathering of influential intellectuals in London, where he was the master of conversation, an ability in which he excelled like nobody else. In 1764, his long acquaintance with Mr and Mrs Thrale (later Mrs. Piozzi) began. He almost became a member of their family. These were also the years of his most mature literary achievements – the Shakespeare edition (1765) and the *Lives of the English poets* (1779–1781).

On June 18, 1746, five booksellers from London contracted with Samuel Johnson to produce a dictionary of English within three years for the sum of 1575 pounds. On April 30 of the same year, Johnson wrote his plan for the dictionary, and he sent a slightly altered version to the Earl of Chesterfield one year later. His hopes of patronage and financial assistance, however, came to nothing. Later, he complained bitterly about this. For nine years, Johnson and six amanuenses worked on the project. Its first volume appeared on April 15, 1755. The two volumes together contained 2300 folio pages with 40 000 entries in total. Each of them consists of the lemma, the determination of word class, in the case of a derived word its stem, etymological references, the definition of meaning(s) by synonym or paraphrase, and quotations from 'good authors.' The amanuenses culled the lemmata and the quotations from books, while Johnson himself provided the definitions and selected the quotations. Both are rightly regarded as praiseworthy and as showing his true genius for language. The dictionary is introduced by a grammar, a history of the language, and a preface. In this preface, Samuel Johnson explained the achievements of his work and admitted that many of his initial aims could not be accomplished. It is a moving testament to his intellectual sincerity.

Samuel Johnson praised the copiousness, energy, variety, and purity of the English language, as used by 'good' authors from Spencer to the Restoration of 1660, but he criticized a lack of order and rules since then. For him, this was the consequence of ignorance, vulgarity, and caprice. He wished to counter these with the ideals of classical rhetoric – brevity, congruity, and perspicuity. It is in this ambition of fixing language use in the sense of these ideals that he admitted having failed. His idealist belief in the necessity of a linguistic norm was thus challenged by the realism of the evaluation of his own work.

The *Dictionary of the English language* was the unquestioned lexicographical authority for language use in England until 1887, when the first volume of what was later to become the *Oxford English dictionary* appeared. It was taken by many as a substitute for a centralized linguistic academy such as the Italians and the French had (and have).

See also: Academies: Dictionaries and Standards; English, Later Modern (ca. 1700–1900); English Lexicography; Rhetoric: History; Synonymy.

Bibliography

Green J (1997). *Chasing the sun. Dictionary makers and the dictionaries they made.* London: Pimlico.

Hüllen W (2004). *The history of Roget's Thesaurus. Origins, development, and design.* Oxford: Oxford University Press. 170–197.

Reddick A (1990). *The making of Johnson's dictionary 1746–1773.* Cambridge: Cambridge University Press.

Sledd J H & Kolb G J (1955). *Dr Johnson's dictionary.* Chicago: Chicago University Press.

Starnes D T & Noyes G E (1991). Ed. by G Stein. *The English dictionary from Cawdrey to Johnson 1604–1755.* Benjamins: Amsterdam. (Orig. 1946.)

Johnston Atoll: Language Situation

Editorial Team

Johnston Atoll, located in the central Pacific Ocean between the Hawaiian and Marshall Islands, is one of the most isolated places in the world. Of volcanic formation, Johnston Atoll is a broad, shallow platform of about 50 square miles with four islands and a marginal, emergent reef on its northwestern side. In 1858, both the United States and the Kingdom of Hawaii annexed the uninhabited Atoll, which eventually became an American possession. Initially, the atoll was exploited for its guano deposits; subsequently, it became a designated wildlife refuge, but this last incarnation did not last long. In the 1930s, the Atoll passed into the hands of the military. Johnston Atoll has been chosen for test launchings of nuclear missiles and it contains a landfill of plutonium-contaminated waste; more than 4 million pounds of chemical weapons have been destroyed on Johnston since 1990. The military has been gradually decreasing its activity and presence on the Atoll, and as of January 2004, Johnston's population had fallen to about 200 Air Force and civilian personnel, all English-speaking.

Jones, Daniel (1881–1967)

B S Collins, Leiden University, Leiden, the Netherlands
I M Mees, Copenhagen Business School, Copenhagen, Denmark

Daniel Jones, the leading British phonetician of the 20th century, was born in London on September 12, 1881, the son of a successful London barrister. He was educated at Ludgrove preparatory school, Radley College, and University College School before going on to read mathematics at Cambridge. Jones graduated with a B.A. in 1903 (converted by payment to M.A. in 1907) but while still a student discovered that he possessed remarkable linguistic talents. Out of interest, he began attending language courses during his vacations, and it was on one such occasion at Marburg, Germany, in 1900 that he was to encounter the renowned language teacher and skilled practical phonetician William Tilly. Jones later explained that this was his introduction to phonetics, and he held Tilly in highest regard thereafter (Collins and Mees, 1999: 12–17).

After graduation, Jones was reluctantly persuaded by his father to take up law. Nevertheless, by pleading ill health, Jones managed to interrupt his legal training in 1905 in order to spend a year in Paris studying under the eminent linguist Paul Passy, whom he soon got to know as a personal friend, and whose niece he married in 1911. Passy acted as guide and mentor to Jones in these early years, encouraging him to take the International Phonetic Association (IPA) examination and to make phonetics his career. Jones performed brilliantly for the IPA certificate, which turned out to be the only linguistic qualification he was ever to obtain (apart from honorary doctorates, Dr. Phil., Zürich, 1936; LL.D., Edinburgh, 1958). He did, however, take private lessons in practical ear-training from Henry Sweet, whom he was later to regard as another strong formative influence.

In 1907, Jones was called to the bar, but never practiced as a barrister, having in the meantime obtained a part-time post teaching phonetics at University College London – then in its infancy as an academic discipline in Britain. The new subject proved popular with students, and eventually, in 1912, Jones was able to set up his own phonetics laboratory and was appointed by London University to be head of the first British university department of phonetics. He was promoted to Reader in 1914 and awarded a chair in 1921 – a post in which he remained until 1949. Upon retirement, he was designated Professor Emeritus.

The year 1912 was also coincidentally when Bernard Shaw completed his first drafts of *Pygmalion,* a play mainly concerned with phonetic matters, and Jones was called upon to provide him with help and advice on technical points. Contrary to previous belief, it now seems likely that it was Daniel Jones, rather than Henry Sweet, who was the chief inspiration for Shaw's fictional phonetician hero Henry Higgins. To avoid any suspicion that Higgins was based on a living phonetician, and thus prevent a possible libel suit, Shaw appears to have attempted to fog the issue by inserting into the Preface to the first published edition of the play (1916) deliberately distracting remarks about the recently deceased Sweet (Collins and Mees, 1999: 97–103).

Johnston Atoll: Language Situation

Editorial Team

Johnston Atoll, located in the central Pacific Ocean between the Hawaiian and Marshall Islands, is one of the most isolated places in the world. Of volcanic formation, Johnston Atoll is a broad, shallow platform of about 50 square miles with four islands and a marginal, emergent reef on its northwestern side. In 1858, both the United States and the Kingdom of Hawaii annexed the uninhabited Atoll, which eventually became an American possession. Initially, the atoll was exploited for its guano deposits; subsequently, it became a designated wildlife refuge, but this last incarnation did not last long. In the 1930s, the Atoll passed into the hands of the military. Johnston Atoll has been chosen for test launchings of nuclear missiles and it contains a landfill of plutonium-contaminated waste; more than 4 million pounds of chemical weapons have been destroyed on Johnston since 1990. The military has been gradually decreasing its activity and presence on the Atoll, and as of January 2004, Johnston's population had fallen to about 200 Air Force and civilian personnel, all English-speaking.

Jones, Daniel (1881–1967)

B S Collins, Leiden University, Leiden, the Netherlands
I M Mees, Copenhagen Business School, Copenhagen, Denmark

Daniel Jones, the leading British phonetician of the 20th century, was born in London on September 12, 1881, the son of a successful London barrister. He was educated at Ludgrove preparatory school, Radley College, and University College School before going on to read mathematics at Cambridge. Jones graduated with a B.A. in 1903 (converted by payment to M.A. in 1907) but while still a student discovered that he possessed remarkable linguistic talents. Out of interest, he began attending language courses during his vacations, and it was on one such occasion at Marburg, Germany, in 1900 that he was to encounter the renowned language teacher and skilled practical phonetician William Tilly. Jones later explained that this was his introduction to phonetics, and he held Tilly in highest regard thereafter (Collins and Mees, 1999: 12–17).

After graduation, Jones was reluctantly persuaded by his father to take up law. Nevertheless, by pleading ill health, Jones managed to interrupt his legal training in 1905 in order to spend a year in Paris studying under the eminent linguist Paul Passy, whom he soon got to know as a personal friend, and whose niece he married in 1911. Passy acted as guide and mentor to Jones in these early years, encouraging him to take the International Phonetic Association (IPA) examination and to make phonetics his career. Jones performed brilliantly for the IPA certificate, which turned out to be the only linguistic qualification he was ever to obtain (apart from honorary doctorates, Dr. Phil., Zürich, 1936; LL.D., Edinburgh, 1958). He did, however, take private lessons in practical ear-training from Henry Sweet, whom he was later to regard as another strong formative influence.

In 1907, Jones was called to the bar, but never practiced as a barrister, having in the meantime obtained a part-time post teaching phonetics at University College London – then in its infancy as an academic discipline in Britain. The new subject proved popular with students, and eventually, in 1912, Jones was able to set up his own phonetics laboratory and was appointed by London University to be head of the first British university department of phonetics. He was promoted to Reader in 1914 and awarded a chair in 1921 – a post in which he remained until 1949. Upon retirement, he was designated Professor Emeritus.

The year 1912 was also coincidentally when Bernard Shaw completed his first drafts of *Pygmalion,* a play mainly concerned with phonetic matters, and Jones was called upon to provide him with help and advice on technical points. Contrary to previous belief, it now seems likely that it was Daniel Jones, rather than Henry Sweet, who was the chief inspiration for Shaw's fictional phonetician hero Henry Higgins. To avoid any suspicion that Higgins was based on a living phonetician, and thus prevent a possible libel suit, Shaw appears to have attempted to fog the issue by inserting into the Preface to the first published edition of the play (1916) deliberately distracting remarks about the recently deceased Sweet (Collins and Mees, 1999: 97–103).

In this initial period of his career, Jones produced a flood of publications, his first commercial success being the compact *Pronunciation of English* (1909). The more ambitious *Outline of English phonetics* (1918), aimed at the nonnative learner, was much more comprehensive and can claim to be the earliest reliable full account of what Jones was later to term 'Received Pronunciation' (RP). Its descriptive framework penetrated linguistics worldwide, influencing Bloomfield and the post-Bloomfieldians of the American Structuralist School. In 1913, Jones co-edited (with Hermann Michaelis) the *Phonetic dictionary of the English language*, the first serious attempt at a dedicated pronunciation dictionary for English in over 100 years. It was soon overtaken by the success of Jones's own *English pronouncing dictionary* (1917a). The *EPD* (as it is always known) rapidly became the recognized authority on RP usage and (in revised form) remains in print to the present day.

Jones pioneered research on non-European languages, producing concise but significant early studies, notably on Cantonese (Jones and Woo, 1912), Tswana (Jones and Plaatje, 1916), and Sinhalese (Sinhala) (Perera and Jones, 1919). Jones's truly groundbreaking research on tone languages is now once again beginning to get the recognition it deserves. His notable description of Tswana contains the first systematic analysis of a register tone language taking in downstep (Jones and Plaatje, 1916: xxv–xxxi). Jones (1917b) expanded on this research; a later innovation was the raised exclamation mark (ꜝ) commonly used for the downstep feature (Jones, 1928: 3). Jones's work on African languages was subsequently continued and expanded by several of his London colleagues and pupils, including Lilias Armstrong, Ida Ward, Clement Doke, Eugénie Henderson, A. N. Tucker, and Jack Carnochan.

Jones did not neglect European languages. The *Pronunciation of Russian* (co-authored with Michael Trofimov), for a long period the only reliable phonetic description in English of the sound system of that language, was published in 1923. At the very end of his life Jones was to rewrite this book together with Dennis Ward, and under the new title *Phonetics of Russian* it appeared in 1969. In 1927, Jones produced (with Emile Stéphan) *Colloquial French,* issued with what at the time was considered a startling innovation – no fewer than 15 gramophone discs containing approximately 90 minutes of spoken material. This proved to be a most successful pioneering example of the audio-lingual method and can be regarded as a milestone in applied linguistics and the history of language teaching.

Jones's (1917c) article, published in *Nature*, is the first detailed exposition of his influential theory of cardinal vowels. Based on ideas derived in part from Paul Passy but also showing the influence of Alexander Melville Bell, Henry Sweet, and the British phonetician Richard Lloyd, it provided a simple dual-parameter model of vowel description based on tongue-arch height and lip-shape. Despite the lack of adequate supporting radiographic evidence, Jones subsequently elaborated the theory, arriving eventually at a system of eight rounded and eight unrounded front and back vowels at supposedly articulatorily equidistant intervals, plus two close central vowels (Collins and Mees, 1995). Jones produced three disc recordings of the cardinal vowels (1917d, 1943, 1956), the last version being nowadays the best known. The cardinal vowel system – although today in some disrepute in the light of contradictory instrumental data – has been widely used since its inception, and much present-day vowel description still derives ultimately from Jones's model (cf. Butcher, 1982). It is noteworthy that a diagram based almost totally on Jones's vowel scheme still appears on the current (1996) IPA chart (see Figure 1).

Jones was the first English writer (from 1917 onward) to use and promulgate the word 'phoneme' in its current sense, borrowing the term from the followers of Baudouin de Courtenay and interpreting it

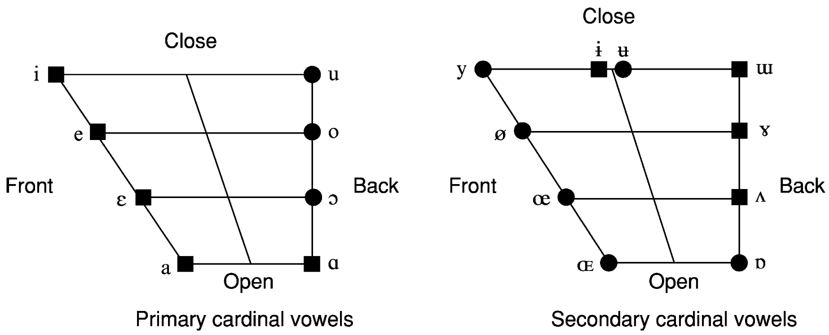


Figure 1 Jones's system of primary and secondary cardinal vowels in updated form. Lip shape is indicated by squares (spread-neutral) and circles (rounded).

as a 'family of sounds' related by physical criteria. As a result, the phoneme principle became firmly established in British linguistics well before the advent of either the Prague School or American structuralism, as Jones himself points out in his *History and meaning of the term 'phoneme'* (1957). After Jones suffered a nervous breakdown in 1920, his dynamic productivity diminished somewhat, but he went on in his later years to write numerous articles and produce important revisions of his major works. His main effort was directed to writing *The phoneme* (1950), which he regarded as his life's work. On publication, even though the large collection of reliable language data was much admired, Jones's static theoretical position – essentially unchanged from 1917 – disappointed contemporary critics.

Jones was, together with Paul Passy, the main force behind the IPA (Assistant Secretary 1906–1927; Secretary 1927–1950; President 1950–1967). He was also editor of its influential journal *Le Maître Phonétique* and from 1906 onward played a leading role in spreading the IPA alphabet worldwide. Jones was an active supporter of English spelling reform and also helped devise schemes for new alphabets for African and Indian languages. In 1926, he became a founding member of the BBC Advisory Committee on Pronunciation and from 1942 to 1967 was chief pronunciation advisor to the BBC. In this role, he was to a degree responsible for reinforcing the notion of 'BBC English' as a standard, but Jones himself was from early on in his career advanced in his libertarian views on pronunciation. Although he revived the Victorian term 'received pronunciation' and the abbreviation RP in the (1926) third edition of the *EPD*, he deprecated any attempt to prescribe a standard pronunciation model for native speakers. He was a keen observer of linguistic change, and was the first to note the spread of glottalization in British English and the increasing influence of popular London speech on RP, thus foreseeing the advance of what is nowadays termed 'estuary English' (Collins and Mees, 2001).

Although now somewhat dated, Jones's works (many of which have recently been re-issued together with hitherto unpublished notes, lectures, and correspondence; Collins and Mees, 2003) remain outstanding both for their detailed phonetic observation and their lucidity. They established him as the best-known phonetician of his generation, and his ideas and methods, propagated by colleagues and pupils forming what Jones (1948) himself termed the London School of Phonetics, have continued to influence many areas of phonetics and pronunciation teaching. (Note that the term London School was later, confusingly, applied to a totally different group of linguists, namely J. R. Firth and his followers.)

Jones can fairly be placed among the leading figures in 20th century linguistics and, together with Henry Sweet, can be considered the greatest of all British phoneticians.

Despite crippling illness, Jones remained working on his publications even in his last years. He died on December 4, 1967, at his home in Gerrards Cross, England.

See also: Armstrong, Lilias Eveline (1882–1937); Bell, Alexander Melville (1819–1905); Doke, Clement Martyn (1893–1980); Phoneme; Phonetic Classification; Phonetic Pedagogy; Phonetic Transcription: Analysis; Shaw, George Bernard (1856–1950); Sweet, Henry (1845–1912); Tilly, William (1860–1935); Ward, Ida Caroline (1880–1950).

Bibliography

Butcher A (1982). 'Cardinal vowels and other problems.' In Crystal D (ed.) *Linguistic controversies.* London: Arnold. 50–72.

Collins B & Mees I M (1995). 'Daniel Jones, Paul Passy and the development of the Cardinal Vowel system.' *Historiographia Linguistica 22,* 197–216.

Collins B & Mees I M (1999). *The real Professor Higgins: the life and career of Daniel Jones.* Berlin: Mouton de Gruyter.

Collins B & Mees I M (2001). 'Daniel Jones, prescriptivist, R.(I.)P.' *English Studies 82(1),* 66–73.

Collins B & Mees I M (eds.) (2003). *Daniel Jones: selected works* (8 vols). London: Routledge.

Fischer-Jørgensen E (1975). *Trends in phonological theory: a historical introduction.* Copenhagen: Akademisk Forlag.

Gimson A C (1968). 'Daniel Jones' (obituary). *Le Maître Phonétique 27(3),* 18–22.

Jones D (1909). *The pronunciation of English.* Cambridge: Cambridge University Press (3rd rewritten edn., 1950; 4th enlarged edn., 1956).

Jones D (1917a). *An English pronouncing dictionary.* London: Dent (3rd edn., 1926).

Jones D (1917b). 'The phonetic structure of the Sechuana language.' *Transactions of the Philological Society 1917–20*, 99–106.

Jones D (1917c). 'Experimental phonetics and its utility to the linguist.' *Nature 100, 96*–98.

Jones D (1917d). *The eight cardinal vowels.* Disc recordings B 804. The Gramophone Company (His Master's Voice).

Jones D (1918). *An outline of English phonetics.* Leipzig: Teubner (9th edn., Cambridge: Cambridge University Press, 1960).

Jones D (1928). *The tones of Sechuana nouns.* London: International Institute of African Languages and Cultures.

Jones D (1943). *The cardinal vowels.* Disc recordings DAJO 1/2. Linguaphone Institute.

Jones D (1948). 'The London School of phonetics.' *Zeitschrift für Phonetik 2(3/4),* 127–135.

Jones D (1950). *The phoneme: its nature and use.* Cambridge: Heffer (3rd edn., 1967).
Jones D (1956). *The cardinal vowels.* Disc recordings ENG 252–255, London & New York: Linguaphone Institute.
Jones D (1957). *The history and meaning of the term 'phoneme.'* London: International Phonetic Association. (Also rpt. as an appendix to 3rd edn. (1967) of *The phoneme*).
Jones D & Plaatje S T (1916). *A Sechuana reader.* London: London University Press.
Jones D & Ward D (1969). *The phonetics of Russian.* Cambridge: Cambridge University Press.
Jones D & Woo K T (1912). *A Cantonese phonetic reader.* London: University of London Press.
Michaelis H & Jones D (1913). *A phonetic dictionary of the English language.* Hanover and Berlin: Carl Meyer (Gustav Prior).
Perera H S & Jones D (1919). *A colloquial Sinhalese reader.* Manchester: Manchester University Press.
Shaw G B (1916). *Pygmalion.* London: Constable.
Stéphan E M & Jones D (1927). *Colloquial French.* London: The Gramophone Company.
Trofimov M & Jones D (1923). *The pronunciation of Russian.* Cambridge: Cambridge University Press.

Jones, William, Sir (1746–1794)

M Pierce, University of Michigan, Ann Arbor, MI, USA

Although the law was his profession, Sir William Jones is most famous for his contributions to linguistics, specifically for his 1786 observation about possible genetic relationships between Sanskrit and certain other languages, which formed the basis of scientific historical/comparative linguistics. A spectacular polyglot (sometimes credited with knowledge of over 20 languages) born in 1746, Sir William showed his interest in languages at an early age, thoroughly mastering Latin, Greek, French, and Italian while a schoolboy at Harrow, and at the same time beginning to study Hebrew and Arabic. In 1764 he enrolled at Oxford, where he continued his studies of languages, adding Portuguese, Spanish, and Persian to his repertoire. In 1766 he became a Fellow of University College, Oxford; in 1772 a Fellow of the Royal Society; and in 1773 a member of the Literary Club headed by Samuel Johnson. In 1768 he received a B.A. from Oxford and an M.A. from the same institution in 1773. Financial pressures eventually caused him to switch his field of study to law, and in 1774 he began practicing law in London. In March 1783 he was appointed to a judgeship in Calcutta and was knighted at the same time. (Edgerton [1946] notes that Jones had been a candidate for this position since 1778, and that his appointment had been delayed because of his open opposition to British action in the American Revolution.) Shortly after arriving in India, Jones founded the Asiatick Society (later renamed the Asiatic Society of Bengal) and presided over the society until his death. Life in India unfortunately adversely affected his health, and Jones died before reaching his forty-eighth birthday in 1794.

Although Jones was not able to devote himself fully to linguistics because of the heavy commitments of his judgeship, and despite his early death, Jones produced a substantial amount of linguistic scholarship (to say nothing of his writings on various legal topics, as well as on other subjects of interest to him, like chess). Jones's most famous work is certainly his lecture, originally presented to the Asiatick Society in February 1786, in which he advanced the theory that Sanskrit was certainly related to Latin and Greek, and possibly to the Celtic, Germanic, and Iranian languages as well. Most importantly, he suggested that this relationship was genetic (i.e., that the relevant languages were descended from a common ancestor). This proposal remains the basis of comparative linguistics.

Even if this had been Jones's only contribution to linguistics, his place in the history of the field would have been secure. However, this was not the only important piece of linguistic scholarship produced by Jones. His interest in Persian had led him to write a Persian grammar, originally published in 1771. Although this work was certainly flawed (see Cannon [1958] for extensive critical discussion), it also filled a gap and remains a testament to Jones's ability (all the more so given that he had apparently taught himself the language and that he was not yet 25 years old when it was published). Jones was also well-known for his translations of Indic and Iranian texts; his first effort in this area, a translation of a Persian manuscript on the life of Nadir Shah into French, was undertaken at the request of the king of Denmark, who owned the manuscript in question, and was published in 1770. The same year he published a book-length study of Persian poetry (Jones, 1770), as well as translations of some of Hafiz' odes into French. Possibly even more important than his

Jones D (1950). *The phoneme: its nature and use.* Cambridge: Heffer (3rd edn., 1967).
Jones D (1956). *The cardinal vowels.* Disc recordings ENG 252–255, London & New York: Linguaphone Institute.
Jones D (1957). *The history and meaning of the term 'phoneme.'* London: International Phonetic Association. (Also rpt. as an appendix to 3rd edn. (1967) of *The phoneme*).
Jones D & Plaatje S T (1916). *A Sechuana reader.* London: London University Press.
Jones D & Ward D (1969). *The phonetics of Russian.* Cambridge: Cambridge University Press.
Jones D & Woo K T (1912). *A Cantonese phonetic reader.* London: University of London Press.
Michaelis H & Jones D (1913). *A phonetic dictionary of the English language.* Hanover and Berlin: Carl Meyer (Gustav Prior).
Perera H S & Jones D (1919). *A colloquial Sinhalese reader.* Manchester: Manchester University Press.
Shaw G B (1916). *Pygmalion.* London: Constable.
Stéphan E M & Jones D (1927). *Colloquial French.* London: The Gramophone Company.
Trofimov M & Jones D (1923). *The pronunciation of Russian.* Cambridge: Cambridge University Press.

Jones, William, Sir (1746–1794)

M Pierce, University of Michigan, Ann Arbor, MI, USA

Although the law was his profession, Sir William Jones is most famous for his contributions to linguistics, specifically for his 1786 observation about possible genetic relationships between Sanskrit and certain other languages, which formed the basis of scientific historical/comparative linguistics. A spectacular polyglot (sometimes credited with knowledge of over 20 languages) born in 1746, Sir William showed his interest in languages at an early age, thoroughly mastering Latin, Greek, French, and Italian while a schoolboy at Harrow, and at the same time beginning to study Hebrew and Arabic. In 1764 he enrolled at Oxford, where he continued his studies of languages, adding Portuguese, Spanish, and Persian to his repertoire. In 1766 he became a Fellow of University College, Oxford; in 1772 a Fellow of the Royal Society; and in 1773 a member of the Literary Club headed by Samuel Johnson. In 1768 he received a B.A. from Oxford and an M.A. from the same institution in 1773. Financial pressures eventually caused him to switch his field of study to law, and in 1774 he began practicing law in London. In March 1783 he was appointed to a judgeship in Calcutta and was knighted at the same time. (Edgerton [1946] notes that Jones had been a candidate for this position since 1778, and that his appointment had been delayed because of his open opposition to British action in the American Revolution.) Shortly after arriving in India, Jones founded the Asiatick Society (later renamed the Asiatic Society of Bengal) and presided over the society until his death. Life in India unfortunately adversely affected his health, and Jones died before reaching his forty-eighth birthday in 1794.

Although Jones was not able to devote himself fully to linguistics because of the heavy commitments of his judgeship, and despite his early death, Jones produced a substantial amount of linguistic scholarship (to say nothing of his writings on various legal topics, as well as on other subjects of interest to him, like chess). Jones's most famous work is certainly his lecture, originally presented to the Asiatick Society in February 1786, in which he advanced the theory that Sanskrit was certainly related to Latin and Greek, and possibly to the Celtic, Germanic, and Iranian languages as well. Most importantly, he suggested that this relationship was genetic (i.e., that the relevant languages were descended from a common ancestor). This proposal remains the basis of comparative linguistics.

Even if this had been Jones's only contribution to linguistics, his place in the history of the field would have been secure. However, this was not the only important piece of linguistic scholarship produced by Jones. His interest in Persian had led him to write a Persian grammar, originally published in 1771. Although this work was certainly flawed (see Cannon [1958] for extensive critical discussion), it also filled a gap and remains a testament to Jones's ability (all the more so given that he had apparently taught himself the language and that he was not yet 25 years old when it was published). Jones was also well-known for his translations of Indic and Iranian texts; his first effort in this area, a translation of a Persian manuscript on the life of Nadir Shah into French, was undertaken at the request of the king of Denmark, who owned the manuscript in question, and was published in 1770. The same year he published a book-length study of Persian poetry (Jones, 1770), as well as translations of some of Hafiz' odes into French. Possibly even more important than his

translations of Persian poetry, however, are his translations of Sanskrit texts. He was the first, for example, to translate the *Sakuntala* into English; his version was itself later translated into a number of European languages, including a German version that inspired Goethe to mention it in one of his own poems. Jones also translated Chinese texts, albeit somewhat less successfully.

The 1993 publication of his collected works, as well as the 1990 biography by Garland Cannon and the consistent citation of his 1786 lecture in historical linguistics handbooks should ensure that Sir William Jones is not forgotten.

See also: Historical and Comparative Linguistics in the 19th Century; Indo–European Languages; Iranian Languages; Sanskrit.

Bibliography

Cannon G (1958). 'Sir William Jones's Persian Linguistics.' *Journal of the American Oriental Society 78,* 262–273.
Cannon G (1990). *The life and mind of Oriental Jones: Sir William Jones, the father of modern linguistics.* Cambridge: Cambridge University Press.
Edgerton F (1946). 'Sir William Jones: 1746–1794.' *Journal of the American Oriental Society 66,* 230–239.
Jones W (1770). *Traité sur la Poésie Orientale.* London.
Jones W (1771). *A grammar of the Persian language.* London: W and J Richardson, Printer.
Jones W (1993). *The Collected Works of Sir William Jones.* New York: New York University Press.

Jonson, Ben (1572–1637)

S Watanabe, Tokyo, Japan

Better known as a dramatist and poet than a grammarian, Benjamin Jonson (usually called Ben Jonson) learned classical literature at Westminster School, under William Camden, and St. John's College, Cambridge. After joining the English troops in Flanders, he returned to London and began working for the stage both as a player and as a playwright in approximately 1595. In 1598 he killed a fellow actor in a duel. Escaping the gallows by benefit of clergy, he became a Roman Catholic during imprisonment, but abjured 12 years later.

His first extant comedy, *Every man in his humour*, which launched him to instant literary fame, was performed by the Lord Chamberlain's company at the Globe in 1598, with Shakespeare in the cast. Following this work, his representative comedies made him the founder of the tradition of 'comedy of humours' or 'comical satires.' His first extant tragedy, *Sejanus*, a play based on Tiberius and his Captain of the Guards, was ill received by the audience at large, but was greatly admired by cultivated persons for his classical erudition.

The accession of James to the throne in 1603 provided him with opportunities to write a long series of court masques. His classical learning and his unique talent for erudite pageantry commended him to King James and Queen Anne. James granted him a pension, thus making him the *de facto* first Poet Laureate. His poems were published in *The forest* (1616) and *Underwoods* (1640). As a literary critic he was the most potent vindicator of the tradition of classical drama and writer in that tradition in sharp contrast to the romantic drama of the preceding Elizabethan Age. His *Timber, or discoveries* (1640) is a landmark achievement in the history of literary criticism. Called the 'rare Ben Jonson,' he was buried in Westminster Abbey with a portrait bust put up on the south wall of Poets' Corner.

Jonson's substantial knowledge of classical literature led him to regard the *status quo* of English as an unruled language. This recognition of his made him one of the very few who wrote a grammar book for English before the 18th century. The relative freedom of grammatical rules in Elizabethan English began to move toward an increasingly stricter code. His *English grammar* is as strongly influenced by Petrus Ramus as many other early English grammarians such as Alexander Gill and Charles Butler were. At the same time, the influences of the traditional Latin grammar books of the Middle Ages such as Alexander de Villa-Dei's *Doctrinal*e are observed in such idiosyncratic technical terms as 'asyndeton' and 'polysyndeton.' The first manuscript of his *English grammar* was destroyed in a fire that ruined his library, one of the richest in England, in 1623. His extant *English grammar* was included posthumously in the two-volume folio edition of his works in 1640.

See also: English, Later Modern (ca. 1700–1900); Ramus, Petrus (1515–1572).

translations of Persian poetry, however, are his translations of Sanskrit texts. He was the first, for example, to translate the *Sakuntala* into English; his version was itself later translated into a number of European languages, including a German version that inspired Goethe to mention it in one of his own poems. Jones also translated Chinese texts, albeit somewhat less successfully.

The 1993 publication of his collected works, as well as the 1990 biography by Garland Cannon and the consistent citation of his 1786 lecture in historical linguistics handbooks should ensure that Sir William Jones is not forgotten.

See also: Historical and Comparative Linguistics in the 19th Century; Indo–European Languages; Iranian Languages; Sanskrit.

Bibliography

Cannon G (1958). 'Sir William Jones's Persian Linguistics.' *Journal of the American Oriental Society 78,* 262–273.
Cannon G (1990). *The life and mind of Oriental Jones: Sir William Jones, the father of modern linguistics.* Cambridge: Cambridge University Press.
Edgerton F (1946). 'Sir William Jones: 1746–1794.' *Journal of the American Oriental Society 66,* 230–239.
Jones W (1770). *Traité sur la Poésie Orientale.* London.
Jones W (1771). *A grammar of the Persian language.* London: W and J Richardson, Printer.
Jones W (1993). *The Collected Works of Sir William Jones.* New York: New York University Press.

Jonson, Ben (1572–1637)

S Watanabe, Tokyo, Japan

Better known as a dramatist and poet than a grammarian, Benjamin Jonson (usually called Ben Jonson) learned classical literature at Westminster School, under William Camden, and St. John's College, Cambridge. After joining the English troops in Flanders, he returned to London and began working for the stage both as a player and as a playwright in approximately 1595. In 1598 he killed a fellow actor in a duel. Escaping the gallows by benefit of clergy, he became a Roman Catholic during imprisonment, but abjured 12 years later.

His first extant comedy, *Every man in his humour*, which launched him to instant literary fame, was performed by the Lord Chamberlain's company at the Globe in 1598, with Shakespeare in the cast. Following this work, his representative comedies made him the founder of the tradition of 'comedy of humours' or 'comical satires.' His first extant tragedy, *Sejanus*, a play based on Tiberius and his Captain of the Guards, was ill received by the audience at large, but was greatly admired by cultivated persons for his classical erudition.

The accession of James to the throne in 1603 provided him with opportunities to write a long series of court masques. His classical learning and his unique talent for erudite pageantry commended him to King James and Queen Anne. James granted him a pension, thus making him the *de facto* first Poet Laureate. His poems were published in *The forest* (1616) and *Underwoods* (1640). As a literary critic he was the most potent vindicator of the tradition of classical drama and writer in that tradition in sharp contrast to the romantic drama of the preceding Elizabethan Age. His *Timber, or discoveries* (1640) is a landmark achievement in the history of literary criticism. Called the 'rare Ben Jonson,' he was buried in Westminster Abbey with a portrait bust put up on the south wall of Poets' Corner.

Jonson's substantial knowledge of classical literature led him to regard the *status quo* of English as an unruled language. This recognition of his made him one of the very few who wrote a grammar book for English before the 18th century. The relative freedom of grammatical rules in Elizabethan English began to move toward an increasingly stricter code. His *English grammar* is as strongly influenced by Petrus Ramus as many other early English grammarians such as Alexander Gill and Charles Butler were. At the same time, the influences of the traditional Latin grammar books of the Middle Ages such as Alexander de Villa-Dei's *Doctrinal*e are observed in such idiosyncratic technical terms as 'asyndeton' and 'polysyndeton.' The first manuscript of his *English grammar* was destroyed in a fire that ruined his library, one of the richest in England, in 1623. His extant *English grammar* was included posthumously in the two-volume folio edition of his works in 1640.

See also: English, Later Modern (ca. 1700–1900); Ramus, Petrus (1515–1572).

Bibliography

Funke O (1940). 'Ben Jonson's English grammar.' *Anglia 64,* 117–134.
Herford C H & Simpson P (eds.) (1925–1952). *Collected works of Ben Jonson* (11 vols). Oxford: Clarendon Press.
Watanabe S (1958). *Studien zur Abhängigkeit der frühneuenglischen Grammatik von den mittelalterlichen Lateingrammatiken.* Ph.D. diss., Münster: University of Münster.

Joos, Martin (1907–1978)

J G Fought, Diamond Bar, CA, USA

Martin Joos [džows] was born in Wisconsin. He studied electrical engineering at the University of Wisconsin for a time, but never completed an undergraduate degree. He was one of the field workers on the Linguistic Atlas of New England project in 1931–1932, then completed an A.M in 1935 and a Ph.D. in 1941 at Wisconsin. He was a lecturer in German at the University of Toronto from 1935 to 1942. From 1942 to 1946, Joos worked as a linguist and cryptographer at the Signal Security Agency in Washington and was an associate editor of *Studies in Linguistics*. He became an Associate Professor in the German department at the University of Wisconsin in 1946 and Professor in 1949, chairing the department from 1962 to 1964. From 1967 to 1972, he was again at Toronto, as Professor and Director of its Center for Linguistic Studies. He received Emeritus rank in 1972.

Joos's most influential work as a Germanist was the meticulously edited *Reader* (1951). He was interested in phonetics as early as 1930, and in 1948 produced the first manual of acoustic phonetics that exploited the new technology of the sound spectrograph. His short paper on the historical and phonological implications of English intervocalic apical stops (1942) is still the best statement on this topic. But his strongest influence on descriptive linguistics was exercised through the selection of papers published in 1957 as *Readings in linguistics.* His selection of papers, perceived omissions, and very personal comments on each paper created a furor; the book was a strong seller for many years. In a memoir on the history of the Linguistic Society (published in part in 1986), his personal views and anecdotes are mingled with more conventional historical details.

Joos was a most unconventional man: he seemed interested in everything around him, yet also absorbed in his thoughts to the point of distraction. He was a very effective teacher. Although his irreverence for social and professional conventions (except as data) would frequently and in some cases permanently exasperate his colleagues, he took his students, sophomores and all, very seriously, treating them with courtesy and patience: the *Readings* were prepared for them, above all. His 1972 paper on a single word in *Parzifal* shows something of what his classes were like. I have heard several versions of a quip about him: "If you ask Martin Joos what time it is, he'll tell you how to build a watch." He might well have accepted this as a compliment. And that watch would keep time.

Bibliography

Hill A (1979). 'Martin Joos 1907–78.' *Language 55,* 665–669.
Joos M (1942). 'A phonological dilemma in Canadian English.' *Language 18,* 141–144.
Joos M (1948). *Acoustic phonetics. Monograph 23.* Linguistic Society of America. Baltimore: Waverly Press.
Joos M (1951). *Middle High German courtly reader.* Madison: University of Wisconsin Press.
Joos M (1972). 'Semantic axiom number one.' *Language 48,* 257–265.
Joos M (1986). *Notes on the development of the Linguistic Society of America, 1924 to 1950.* Ithaca, NY: Linguistica.

Bibliography

Funke O (1940). 'Ben Jonson's English grammar.' *Anglia 64,* 117–134.

Herford C H & Simpson P (eds.) (1925–1952). *Collected works of Ben Jonson* (11 vols). Oxford: Clarendon Press.

Watanabe S (1958). *Studien zur Abhängigkeit der frühneuenglischen Grammatik von den mittelalterlichen Lateingrammatiken.* Ph.D. diss., Münster: University of Münster.

Joos, Martin (1907–1978)

J G Fought, Diamond Bar, CA, USA

Martin Joos [džows] was born in Wisconsin. He studied electrical engineering at the University of Wisconsin for a time, but never completed an undergraduate degree. He was one of the field workers on the Linguistic Atlas of New England project in 1931–1932, then completed an A.M in 1935 and a Ph.D. in 1941 at Wisconsin. He was a lecturer in German at the University of Toronto from 1935 to 1942. From 1942 to 1946, Joos worked as a linguist and cryptographer at the Signal Security Agency in Washington and was an associate editor of *Studies in Linguistics*. He became an Associate Professor in the German department at the University of Wisconsin in 1946 and Professor in 1949, chairing the department from 1962 to 1964. From 1967 to 1972, he was again at Toronto, as Professor and Director of its Center for Linguistic Studies. He received Emeritus rank in 1972.

Joos's most influential work as a Germanist was the meticulously edited *Reader* (1951). He was interested in phonetics as early as 1930, and in 1948 produced the first manual of acoustic phonetics that exploited the new technology of the sound spectrograph. His short paper on the historical and phonological implications of English intervocalic apical stops (1942) is still the best statement on this topic. But his strongest influence on descriptive linguistics was exercised through the selection of papers published in 1957 as *Readings in linguistics.* His selection of papers, perceived omissions, and very personal comments on each paper created a furor; the book was a strong seller for many years. In a memoir on the history of the Linguistic Society (published in part in 1986), his personal views and anecdotes are mingled with more conventional historical details.

Joos was a most unconventional man: he seemed interested in everything around him, yet also absorbed in his thoughts to the point of distraction. He was a very effective teacher. Although his irreverence for social and professional conventions (except as data) would frequently and in some cases permanently exasperate his colleagues, he took his students, sophomores and all, very seriously, treating them with courtesy and patience: the *Readings* were prepared for them, above all. His 1972 paper on a single word in *Parzifal* shows something of what his classes were like. I have heard several versions of a quip about him: "If you ask Martin Joos what time it is, he'll tell you how to build a watch." He might well have accepted this as a compliment. And that watch would keep time.

Bibliography

Hill A (1979). 'Martin Joos 1907–78.' *Language 55,* 665–669.

Joos M (1942). 'A phonological dilemma in Canadian English.' *Language 18,* 141–144.

Joos M (1948). *Acoustic phonetics. Monograph 23.* Linguistic Society of America. Baltimore: Waverly Press.

Joos M (1951). *Middle High German courtly reader.* Madison: University of Wisconsin Press.

Joos M (1972). 'Semantic axiom number one.' *Language 48,* 257–265.

Joos M (1986). *Notes on the development of the Linguistic Society of America, 1924 to 1950.* Ithaca, NY: Linguistica.

Jordan: Language Situation

E Al-Wer, University of Essex, Colchester, UK

Jordan is predominantly an Arabic-speaking country. It occupies an area of 92 000 km^2 and borders Syria (on the north), Iraq and Saudi Arabia (eastern border), the Gulf of Aqaba (on the south), and Palestine and Israel (western border). Jordan became an independent political entity in 1923, first as an emirate under British protection, then as an independent kingdom, since 1946. Its population in 1946 was no more than half a million; Jordan now has 5.4 million people. In addition to a very high birth rate, the growth in Jordan's population has resulted from the displacement of Palestinians from Israel, the West Bank, the Gaza Strip, and Kuwait (during the Gulf War). The country is also host to hundreds of thousands of workers from Egypt, Sri Lanka, the Philippines, Pakistan, and Iraq (since 1990).

English, the major foreign language, is taught in all state and private schools as a compulsory subject and is used as the medium of instruction in the faculties of science and medicine in higher education. It is also widely used in the business and banking sectors, and as a lingua franca in communications with non-Arab residents. French is considered the second foreign language, and is taught as such in many private schools, but is not ordinarily used outside the classroom. Jordanian Arabic borrows extensively from English and contains a number of borrowings from French.

The most important ethnic and linguistic minority group is Circassian (estimated to be 60 000 people). The Circassian language (a North Caucasian language of the western group) is undergoing a shift in favor of Arabic. The second most important group is the Chechen-Ingush (an estimated 10 000 people); members of this group are generally bilingual in Chechen-Ingush (a North Caucasian language, of the eastern group) and Arabic. The Circassians and the Chechens in Jordan are descendents of the 19th- and 20th-century emigrants from North Caucasia. Their languages are mutually unintelligible (for details on both communities, see Al-Wer, 1999). Another minority language, Armenian, is spoken by 4000 people (see Al-Khatib, 2001). Domari is spoken by some 35 000 people. There is also a small community of Turkomans, speakers of a variety of Turkish (see Moawwad, 1999). Jordan has a national sign language (Arabic acronym, LIU).

Jordanian Arabic is classified into two major dialect groups. Broadly speaking, the eastern and southern dialects are of the Najdi type and have fewer speakers compared to the northern and northwestern dialects, which are classified as (southern) Levantine (see Cleveland, 1963). Within the Levantine group, the most important distinction is between the Haurani type, exemplified by the dialects of the cities of Ajloun (in the north), and Sult (in the northwest), and the Madani (or urban) type. The Madani dialects are very similar to those spoken in the major cities in the region (e.g., Jerusalem, Damascus, and Beirut). The emerging dialect of Amman, the capital city of Jordan, has a predominance of phonological features that are akin to the urban Levantine group (for details, see Al-Wer, 2002).

See also: Arabic; Variation in Arabic Languages.

Language Maps (Appendix 1): Map 104.

Bibliography

Al-Khatib M (2001). 'Language shift among the Armenians of Jordan.' *International Journal of the Sociology of Language 152,* 153–177.

Al-Wer E (1999). 'Language and identity: The Chechens and the Circassians in Jordan.' *Dirasat, Special Issue* 253–268.

Al-Wer E (2002). 'Jordanian and Palestinian dialects in contact: vowel raising in Amman.' In Jones M & Esch E (eds.) *Language change. The interplay of internal, external and extra-linguistic factors.* Berlin: Mouton de Gruyter. 63–79.

Cleveland R (1963). 'A classification of the Arabic dialects of Jordan.' *Bulletin of the American School of Oriental Research 167,* 56–63.

Moawwad K (1999). 'The linguistic situation of gypsies and Turkomans as ethnic minorities living in Jordan: a sociolinguistic perspective.' MA thesis, University of Yarmouk, Jordan.

Jud, Jakob (1882–1952)

S N Dworkin, University of Michigan, Ann Arbor, MI, USA

Jakob Jud spent his entire university teaching career at the University of Zurich, where he studied under H. Morf and L. Gauchat. He also trained under J. Gilliéron and M. Roques in Paris.

Jud specialized in dialect geography and lexical paleontology, stressing the relationship between words and culture. Jud first came to the attention of his fellow Romanists with a series of early papers based on his Habilitationsschrift (*Sprachgeographische Untersuchungen*), which challenged, on the basis of lexical distribution, several etymologies proposed by Wilhelm Meyer-Lübke. Unlike Gilliéron, who limited his purview to French and to the synchronic data available in his own *Atlas linguistique de la France*, Jud viewed the geographic distribution of lexical items from a dynamic and pan-Romance perspective; excellent examples are the word studies in Jud (1925–1926). He examined lexical variation across the ROMANIA to demonstrate how the study of Romance loans in neighboring Germanic and Celtic languages could throw light on the diffusion of Latin lexical items in the late Roman Empire (Jud 1914, 1946). The close connection between lexical variation and culture can be seen in his essays on French, Italian, and Rhaeto-Romance ecclesiastical terminology (reprinted in Jud, 1973). Most of Jud's lexical studies deal with these languages, with occasional incursions into Spanish.

His most important single undertaking was the *Sprach- und Sachatlas Italiens und der Südschweiz* (1928–1940), prepared in collaboration with his friend and Bern colleague Karl Jaberg. From the beginning, Jud had a deep, abiding interest in Rhaeto-Romance language and culture. He wrote a number of studies to demonstrate the linguistic independence of this language group and to combat the claims made in fascist Italy that Rhaeto-Romance was an extension of Northern Italian varieties. As a result of the growing inaccessibility in the thirties of German journals to foreign specialists, Jud cofounded (with Arnald Steiger) the journal *Vox Romanica*, which continues to be the organ of Swiss Romanists. Sixteen representative studies and a complete bibliography of Jud's publications can be found in Jud (1973). Appreciations of his life and work can be found in Steiger (1952) and Kristol (1997).

See also: Gauchat, Louis (1866–1942); Gilliéron, Jules (1854–1926); Jaberg, Karl (1877–1959); Meyer-Lübke, Wilhelm (1861–1936).

Bibliography

Jaberg K & Jud J (1928–40). *Sprach- und Sachatlas Italiens und der Südschweiz* (8 vols). Zofingen: Ringier and Bern: Stänpfli.

Jud J (1914). 'Probleme der altromanischen Wortgeographie.' *Zeitschrift für romanische Philologie 38,* 1–75.

Jud J (1925–1926). 'Problèmes de géographie linguistique.' *Revue de linguistique romane 1,* 181–236.

Jud J (1946). 'Zur Geschichte der romanischen Reliktwörter in den Alpenmundarten der deutsche Schweiz.' *Vox Romanica 8,* 34–109.

Jud J (1973). *Romanische Sprachgeschichte und Sprachgeographie: Ausgewählte Aufsätze*. Zurich: Atlantis.

Kristol A (1997). 'Jakob Jud (1882–1952).' In Wüest J (ed.) *Les linguistes suisses et la variation linguistique. Actes d'un colloque organisé à l'occasion du centenaire du Séminaire des langues romanes de l'Université de Zurich. Romanica Helvetica 116*. Basel: Francke. 125–136.

Steiger A (1952). 'In memoriam Jakob Jud.' *Vox Romanica 12,* ix–xix.

Judaism

J Neusner, Bard College, Annandale-on-Hudson, NY, USA

The Language of Judaism

Hebrew, an ancient northwest Semitic language, has served as the principal language of Judaism, even after the Jews ceased to speak Hebrew as their everyday language. The Hebrew Scriptures ('Old Testament,' 'written Torah') of ancient Israel were mainly in Hebrew. Brief sections of Ezra and Daniel are in Aramaic. The Scriptures were translated into Greek, Aramaic, and other languages, but in the synagogue were and are declaimed in Hebrew. The great commentaries to Scripture, written by the Rabbinic sages of the first six centuries of the Common Era, were written in Hebrew, and philosophy, poetry, and liturgy favored that language as well. That preference for Hebrew persisted, even though it was explicitly

Jud, Jakob (1882–1952)

S N Dworkin, University of Michigan, Ann Arbor, MI, USA

Jakob Jud spent his entire university teaching career at the University of Zurich, where he studied under H. Morf and L. Gauchat. He also trained under J. Gilliéron and M. Roques in Paris.

Jud specialized in dialect geography and lexical paleontology, stressing the relationship between words and culture. Jud first came to the attention of his fellow Romanists with a series of early papers based on his Habilitationsschrift (*Sprachgeographische Untersuchungen*), which challenged, on the basis of lexical distribution, several etymologies proposed by Wilhelm Meyer-Lübke. Unlike Gilliéron, who limited his purview to French and to the synchronic data available in his own *Atlas linguistique de la France*, Jud viewed the geographic distribution of lexical items from a dynamic and pan-Romance perspective; excellent examples are the word studies in Jud (1925–1926). He examined lexical variation across the ROMANIA to demonstrate how the study of Romance loans in neighboring Germanic and Celtic languages could throw light on the diffusion of Latin lexical items in the late Roman Empire (Jud 1914, 1946). The close connection between lexical variation and culture can be seen in his essays on French, Italian, and Rhaeto-Romance ecclesiastical terminology (reprinted in Jud, 1973). Most of Jud's lexical studies deal with these languages, with occasional incursions into Spanish.

His most important single undertaking was the *Sprach- und Sachatlas Italiens und der Südschweiz* (1928–1940), prepared in collaboration with his friend and Bern colleague Karl Jaberg. From the beginning, Jud had a deep, abiding interest in Rhaeto-Romance language and culture. He wrote a number of studies to demonstrate the linguistic independence of this language group and to combat the claims made in fascist Italy that Rhaeto-Romance was an extension of Northern Italian varieties. As a result of the growing inaccessibility in the thirties of German journals to foreign specialists, Jud cofounded (with Arnald Steiger) the journal *Vox Romanica*, which continues to be the organ of Swiss Romanists. Sixteen representative studies and a complete bibliography of Jud's publications can be found in Jud (1973). Appreciations of his life and work can be found in Steiger (1952) and Kristol (1997).

See also: Gauchat, Louis (1866–1942); Gilliéron, Jules (1854–1926); Jaberg, Karl (1877–1959); Meyer-Lübke, Wilhelm (1861–1936).

Bibliography

Jaberg K & Jud J (1928–40). *Sprach- und Sachatlas Italiens und der Südschweiz* (8 vols). Zofingen: Ringier and Bern: Stänpfli.
Jud J (1914). 'Probleme der altromanischen Wortgeographie.' *Zeitschrift für romanische Philologie 38,* 1–75.
Jud J (1925–1926). 'Problèmes de géographie linguistique.' *Revue de linguistique romane 1,* 181–236.
Jud J (1946). 'Zur Geschichte der romanischen Reliktwörter in den Alpenmundarten der deutsche Schweiz.' *Vox Romanica 8,* 34–109.
Jud J (1973). *Romanische Sprachgeschichte und Sprachgeographie: Ausgewählte Aufsätze.* Zurich: Atlantis.
Kristol A (1997). 'Jakob Jud (1882–1952).' In Wüest J (ed.) *Les linguistes suisses et la variation linguistique. Actes d'un colloque organisé à l'occasion du centenaire du Séminaire des langues romanes de l'Université de Zurich. Romanica Helvetica 116.* Basel: Francke. 125–136.
Steiger A (1952). 'In memoriam Jakob Jud.' *Vox Romanica 12,* ix–xix.

Judaism

J Neusner, Bard College, Annandale-on-Hudson, NY, USA

The Language of Judaism

Hebrew, an ancient northwest Semitic language, has served as the principal language of Judaism, even after the Jews ceased to speak Hebrew as their everyday language. The Hebrew Scriptures ('Old Testament,' 'written Torah') of ancient Israel were mainly in Hebrew. Brief sections of Ezra and Daniel are in Aramaic. The Scriptures were translated into Greek, Aramaic, and other languages, but in the synagogue were and are declaimed in Hebrew. The great commentaries to Scripture, written by the Rabbinic sages of the first six centuries of the Common Era, were written in Hebrew, and philosophy, poetry, and liturgy favored that language as well. That preference for Hebrew persisted, even though it was explicitly

stated that translation of Scripture into Aramaic was allowed, and the use of the vernacular for various purposes made explicit.

The prophet Zephaniah (3:9) speaks of the future, in which "I (God) will change the speech of the peoples to a pure speech," meaning Hebrew. Called 'lashon haqqodesh,' or Holy Language, the Hebrew language is accorded religious status in Judaism. A community with roots deep in the past, diverse and widely scattered, affirms its coherence through privileging a single language, Hebrew. Hebrew is a way for "identifying one's original religious allegiance" (Aaron, 2000a: 271). "By relating to their language as holy, Jews transformed Hebrew into a kind of ritual object, parallel . . . to the Torah-scroll itself . . . part of a religious system" (Aaron, 2000a: 268–9).

Language of Prayer; Languages of Daily Life

Even though Hebrew enjoyed the status of a religious artifact in Judaism, the languages of the Israelites in the Land of Israel in late antiquity were Aramaic and Greek. Leading Rabbinic sages knew Greek and encouraged some of their disciples to use it (Mishnah-tractate Abodah Zarah 3:4). While it was not the language of daily life, Hebrew enjoyed privileged standing in certain liturgical settings, but by no means throughout. But since not everybody knew Hebrew, provision was made for the recitation of obligatory prayers in any language, and these included the recitation of the creed, the Shema ("Hear O Israel, the Lord our God, the Lord is one") and the presentation of the Prayer ("Eighteen benedictions").

> Mishnah-tractate Sotah 7:1–2
>
> 7:1 A. These are said in any language: (1) the pericope of the accused wife (Num. 5:19–22), and (2) the confession of the tithe (Deut. 26:13–15), and (3) the recital of the Shema, (Deut. 6:4–9), and (4) the Prayer, (5) the oath of testimony, and (6) the oath concerning a bailment.
>
> 7:2 A. And these are said (only) in the Holy Language (Hebrew): (1) the verses of the first fruits (Deut. 26:3–10), (2) the rite of removing the shoe (Deut. 25:7), (3) blessings and curses (Deut. 27:15–26), (4) the blessing of the priests (Num. 6:24–26), (5) the blessing of a high priest (on the Day of Atonement), (6) the pericope of the king (Deut. 17:14–20);
>
> B. (7) the pericope of the heifer whose neck is to be broken (Deut. 21:7f.), and (8) (the message of) the anointed for battle when he speaks to the people (Deut. 20:2–7).

What is limited to Hebrew are few items, mostly involving texts supplied by Scripture, as indicated. Here the key is that the recitation must be verbatim, hence not only a fixed text, as in the Shema and the Prayer, but a set text that Scripture prescribes, in so many words, in Hebrew.

Translations of Scripture into Aramaic (called Targumim) and Greek were enunciated in public worship, but Scripture had to be read, first of all, in Hebrew. The following concerns the public declamation, on Purim, of the scroll of Esther, which had to be read in Hebrew, not in translation:

> Mishnah-tractate Megillah 2:1
>
> A. One who reads the Megillah (scroll of Esther) out of (its literary) sequence has not fulfilled his obligation.
>
> B. (1) If he read it from memory, (2) if he read it in translation into any language, he has not fulfilled his obligation; (3) but one may read it to non-speakers of Hebrew (*le'uzot*) in other languages. (4) And a speaker of another language who heard Ashurit (the text read from a scroll written in Hebrew language and in square script) has fulfilled his obligation.

But that is only one aspect of the matter of public worship and communal rites. When it comes to prayer, Hebrew was not privileged. The Shema, the Credo recited in worship, could be said in any language, so too the formal petitions or The Prayer par excellence were said in the vernacular. But in the history of Judaism from antiquity forward, the use of Hebrew in synagogue worship predominated. Home rites, such the liturgy, Haggadah ('Narrative') recited at the Passover Seder, also were developed principally in Hebrew. Prayer other than in Hebrew, while licit, thus was discouraged: "When one petitions for his needs in Aramaic, the ministering angels do not listen to him, for they do not understand Aramaic" (Neusner, 1995; Bavli Shabbat, 112b).

Aramaic was the everyday language of the Rabbis of the Judaic academies, but they produced in Hebrew the law code, the Mishnah, and its supplement, the Tosefta, and commentaries, the Talmud of the Land of Israel and the Talmud of Babylonia. Their Hebrew differed from that of Scripture. But education in the Hebrew of Scripture and of the Rabbinic writings was integral to the religious nurture of the young (see Neusner, 1987; Sifre to Deuteronomy XLVI:I).

The Role of Language Analysis in the Law of Judaism

The legal literature of Judaism is written in Hebrew and in Aramaic. The law codes, the Mishnah, Tosefta, and related collections, are in Hebrew. The analytical work of commenting upon the law is in Aramaic. By their choice of the very language in which they would express what they wished to say on their own account

they differentiated themselves from their antecedents. When it came to citations from prior, non-scriptural authorities, they used one formation of the Hebrew language, specifically, Middle, or Mishnaic, Hebrew; when it came to the conduct of their own analytical process, they used one formation of the Aramaic language, Eastern or Talmudic Aramaic. They never alluded to authoritative facts, they always cited them in so many words; but the indication of citation – in a writing in which the modern sigla of quotation marks and footnotes were simply unavailable – came to expression in the choice of language. In point of fact, the Talmud of Babylonia or Bavli is in one language, not two, and that language is Aramaic. The infrastructure of the document, its entire repertoire of editorial conventions, and sigla are in Aramaic.

When a saying is assigned to a named authority, the saying may be in Hebrew or in Aramaic, and the same named authority may be given sayings in both languages – even within the same sentence. But the editorial and conceptual infrastructure of the document comes to expression only in Aramaic, and when no name is attached to a statement, that statement is always in Aramaic, unless it forms part of a larger, autonomous Hebrew composition, cited by, or parachuted down into, 'the Talmud.' Rightly have the Talmudic masters in the Yeshiva-world hypostatized the Talmud in such language as, "the Gemara says. . ." because the Talmud speaks in a single voice; forms a unitary discourse, beginning, middle, and end; and constitutes one wholly coherent and cogent document, everywhere asking questions drawn from a single determinate and limited repertoire of intellectual initiatives – and always framing those questions, pursuing those inquiries, in Aramaic.

And yet the Talmud also is full of Hebrew. So we must ask where and why framers of this writing utilize the Hebrew language, and when we may expect to find that they speak – rather, 'the Talmud speaks' – in Aramaic. Specifically, what signal is given, what purpose is served by the bi- or multilingualism of the Talmud, what do we know without further ado, when we are given a composition or a component of a composition in Hebrew, and what is the implicit meaning of making a statement in Aramaic? The answer is that the choice of language signals a taxonomic meaning, and language serves as a medium for the classification of discourse, hence, language serves as a medium of taxonomy. In a writing that utilizes two languages, the choice of one over the other conformed to rules of communication and marked what was said as one type of statement rather than another. If we know which language is used, we also know where we stand in the expression of thought, and the very language in which a statement is made therefore forms part of the method of thought and even the message of discourse of the document.

The Role of Language Analysis in the Theology and Exegesis of Judaism

In reading Scripture, Judaism brought to bear a particular theory of how language works (Aaron, 2005: 400). The Rabbinic exegetes of Scripture distinguished their version of Hebrew from that of Scripture; italics signify Aramaic, plain type, Hebrew:

Bavli Hullin 137b

V.5 A. *When Issi bar Hini left (to go to the Land of Israel), R. Yohanan found him teaching his son (our Mishnah at L, with a variation in the plural suffix of the word sheep): "Five sheep (masculine plural suffix,* rhylym)."

B. *He (Yohanan) said to him, "It was taught (in the Mishnah differently, namely using the plural feminine suffix): 'Sheep (rhylwt).'"*

C. *He (Issi) said to him, "That is the way it is written in scripture,* 'Two hundred sheep (rhylym) (Gen. 32:15).'"

D. *He (Yohanan) said to him,* "The Torah has its own language and the sages have their own language. (You should not confuse the two.)"

Their rules of exegesis corresponded with ordinary speech. Thus the argument from analogy (*gezerah shavah*) is a commonplace. Verses that share the same word choice are asked to clarify the meaning of one another. Various particles of speech serve to signal an inclusive meaning of a given concept. Thus, in Aaron's words (Aaron, 2005: 402), "The rules of common logic that normally govern colloquial discourse are transformed by virtue of the Rabbinic conceptualization of the Torah's language. This transformation results from a distinct theory of word meaning, one that allows for exegetical processes that violate what we would otherwise consider the norms of everyday speech. The upshot is that the text of the Torah bears meanings below its surface. When God revealed his words through Moses his prophet, he provided for layers and layers of meaning.

The traits of the Hebrew language, including its alphabet, yielded theological lessons. "The very letters themselves receive their forms by virtue of divine intent" (Aaron, 2005: 407).

Genesis Rabbah I:X.

1. A. ("In the beginning God created" (Gen. 1:1):) R. Jonah in the name of R. Levi: "Why was the world created with (a word beginning with the letter) B?

B. "Just as (in Hebrew) the letter B is closed (at the back and sides but) open in front, so you have no right to

expound concerning what is above or below, before or afterward."

C. Bar Qappara said, "For ask now of the days past which were before you, since the day that God created man upon the earth" (Deut. 4:32).

D. "Concerning the day *after* which days were created, you may expound, but you may not make an exposition concerning what lies before them"

E. "And from one end of the heaven to the other" (Deut. 4:32).

F. "(Concerning that space) you may conduct an investigation, but you may not conduct an investigation concerning what lies beyond those points."

Aaron explains, "Midrash relies on a concept of meaning that violates the rules of everyday language usage, even as it relies upon them to convey its meanings . . . Midrash becomes a vehicle in ordinary language by which to uncover hidden meanings embedded in the extraordinary language of God's revelation" (Aaron, 2005: 410).

Conclusion

Until the beginning of the 20th century, Hebrew served as the language of synagogue worship and other ritual occasions, study of the Torah or divine revelation, and legal and theological discourse. It was the international language of Judaism, superseding the local languages spoken by Jews. It was comparable in function to Latin in Western Christendom. This ritualization of the language is explained by Aaron in the following language (Aaron, 2000a: 283):

> During the Rabbinic period, despite relatively low levels of Hebrew literacy among common Jews, Hebrew began serving as Judaism's language in ritual contexts. The success of this ritualization of language was part of a conceptualization of language that sought to transcend the common level of semantics we normally ascribe to simple utterances . . . With ritualization, meaning could be found not only in the semantic values of the words read but also in the act of recitation itself . . . The whole Rabbinic enterprise is founded upon a meta-semantics of sorts, one that holds that meaning is at once rooted in the words of Torah but not identical to their common connotations."

Hebrew was and is by no means the only language identified with the Jewish group and with Judaism. Yiddish (based on Hebrew and German), Ladino (based on Hebrew and Spanish), Judeo-Arabic, Judeo-Persian, and other local languages have flourished. But when it came to the creation of the Jewish state, the State of Israel, it was Hebrew that was chosen for renewal, and that became the language of the new and renewed culture of the Jewish people.

See also: Aramaic and Syriac; Hebrew, Biblical and Jewish; Hebrew, Israeli.

Bibliography

Aaron D (2000a). 'The doctrine of Hebrew language usage.' In Neusner J & Avery-Peck A J (eds.), 268–287.

Aaron D (2000b). 'Readings in the doctrine of Hebrew language usage.' In Neusner J & Avery-Peck A J (eds.), 202–214.

Aaron D (2005). 'Language and Midrash.' In Neusner J *et al.* (eds.) *Encyclopaedia of Midrash*. Leiden: Brill. 400–411.

Hoffman J M (2004). *In the beginning: a short history of the Hebrew language*. New York: New York University Press.

Neusner J (1987). *Sifré to Deuterotomy. An Analytical Translation*. Atlanta, 1987: Scholars Press for Brown Judiac Studies. (Now) Lanham: University Press of America.

Neusner J (1990). *Language as taxonomy: the rules for using Hebrew and Aramaic in the Babylonian Talmud*. Atlanta: Scholars Press for South Florida Studies in the History of Judaism. (Now) Lanham, MD: University Press of America.

Neusner J (1995). *Talmud of Babylonia, tractate Shabbat*. Lanham, 1995: University Press of America. *The Talmud of Babylonia. An American Translation* by Jacob Neusner.

Neusner J & Avery-Peck A J (eds.) (2000). *The Blackwell companion to Judaism*. Oxford: Blackwell.

Junius (Francisci Filius), Franciscus (1591–1677)

L Kaczmarek, Borgholzhausen, Germany

Franciscus (or Francis) Junius F. F. (called 'filius,' son, to distinguish him from his nephew of the same name) was born on January 29, 1591, in Heidelberg, Germany, the son of the French-born Calvinist theologian François du Jon (Junius the Elder). Gerardus J. Vossius, the renowned humanist who became his brother-in-law, was headmaster of his Latin school and interested him in more advanced studies of the older stages of the Germanic languages. Junius studied the arts and theology (under his uncle, the anti-Remonstrant Gomarus) at

expound concerning what is above or below, before or afterward."

C. Bar Qappara said, "For ask now of the days past which were before you, since the day that God created man upon the earth" (Deut. 4:32).

D. "Concerning the day *after* which days were created, you may expound, but you may not make an exposition concerning what lies before them"

E. "And from one end of the heaven to the other" (Deut. 4:32).

F. "(Concerning that space) you may conduct an investigation, but you may not conduct an investigation concerning what lies beyond those points."

Aaron explains, "Midrash relies on a concept of meaning that violates the rules of everyday language usage, even as it relies upon them to convey its meanings . . . Midrash becomes a vehicle in ordinary language by which to uncover hidden meanings embedded in the extraordinary language of God's revelation" (Aaron, 2005: 410).

Conclusion

Until the beginning of the 20th century, Hebrew served as the language of synagogue worship and other ritual occasions, study of the Torah or divine revelation, and legal and theological discourse. It was the international language of Judaism, superseding the local languages spoken by Jews. It was comparable in function to Latin in Western Christendom. This ritualization of the language is explained by Aaron in the following language (Aaron, 2000a: 283):

> During the Rabbinic period, despite relatively low levels of Hebrew literacy among common Jews, Hebrew began serving as Judaism's language in ritual contexts. The success of this ritualization of language was part of a conceptualization of language that sought to transcend the common level of semantics we normally ascribe to simple utterances . . . With ritualization, meaning could be found not only in the semantic values of the words read but also in the act of recitation itself . . . The whole Rabbinic enterprise is founded upon a meta-semantics of sorts, one that holds that meaning is at once rooted in the words of Torah but not identical to their common connotations."

Hebrew was and is by no means the only language identified with the Jewish group and with Judaism. Yiddish (based on Hebrew and German), Ladino (based on Hebrew and Spanish), Judeo-Arabic, Judeo-Persian, and other local languages have flourished. But when it came to the creation of the Jewish state, the State of Israel, it was Hebrew that was chosen for renewal, and that became the language of the new and renewed culture of the Jewish people.

See also: Aramaic and Syriac; Hebrew, Biblical and Jewish; Hebrew, Israeli.

Bibliography

Aaron D (2000a). 'The doctrine of Hebrew language usage.' In Neusner J & Avery-Peck A J (eds.), 268–287.

Aaron D (2000b). 'Readings in the doctrine of Hebrew language usage.' In Neusner J & Avery-Peck A J (eds.), 202–214.

Aaron D (2005). 'Language and Midrash.' In Neusner J *et al.* (eds.) *Encyclopaedia of Midrash*. Leiden: Brill. 400–411.

Hoffman J M (2004). *In the beginning: a short history of the Hebrew language.* New York: New York University Press.

Neusner J (1987). *Sifré to Deuterotomy. An Analytical Translation.* Atlanta, 1987: Scholars Press for Brown Judiac Studies. (Now) Lanham: University Press of America.

Neusner J (1990). *Language as taxonomy: the rules for using Hebrew and Aramaic in the Babylonian Talmud.* Atlanta: Scholars Press for South Florida Studies in the History of Judaism. (Now) Lanham, MD: University Press of America.

Neusner J (1995). *Talmud of Babylonia, tractate Shabbat.* Lanham, 1995: University Press of America. *The Talmud of Babylonia. An American Translation* by Jacob Neusner.

Neusner J & Avery-Peck A J (eds.) (2000). *The Blackwell companion to Judaism.* Oxford: Blackwell.

Junius (Francisci Filius), Franciscus (1591–1677)

L Kaczmarek, Borgholzhausen, Germany

Franciscus (or Francis) Junius F. F. (called 'filius,' son, to distinguish him from his nephew of the same name) was born on January 29, 1591, in Heidelberg, Germany, the son of the French-born Calvinist theologian François du Jon (Junius the Elder). Gerardus J. Vossius, the renowned humanist who became his brother-in-law, was headmaster of his Latin school and interested him in more advanced studies of the older stages of the Germanic languages. Junius studied the arts and theology (under his uncle, the anti-Remonstrant Gomarus) at

Leiden. For a short time, Junius served as a priest in Hillegersberg but then resigned and went to England, where he lived with the Howard family as a tutor, antiquarian, and art historian for more than 20 years, publishing his *De Pictura Veterum* in 1637.

Back in Holland in 1642, he began his second phase (which lasted until his death in Windsor, England, on November 19, 1677), first as an etymologist interested in roots, and later as a lexicographer developing corpora. He also collected and edited medieval manuscripts and old books, which he heavily annotated. Since about 1644, he studied the Old Germanic languages (which he came to accept as a grouping of its own) and planned an etymological dictionary of Dutch, his *Etymologicum Teutonicum*, a project he never realized. In 1647–1648, he studied the Frisian language in Friesland (see article by P. H. Breuker in Bremmer 1998: 129–157).

Returning to England in 1648, Junius advised William Somner, who worked on a *Dictionarium Saxonico–Latino–Anglicum* (1659). From 1651 to 1674, Junius lived in Holland again, editing and commenting Old High German (*Observationes in Willerami*, 1655, repr. 1992) and Old English texts (*Cædmonis monachi paraphrasis poetica*, 1655, repr. 2000). From 1654 he was able to work on the sixth-century *Codex Argenteus* that his nephew Isaac Vossius had received as a present from Sweden. It contains the Gothic translations of parts of the Gospels by Bishop Wulfila (fourth century). This work resulted in Junius's *Gothicum Glossarium* (1665), the first lexicon of Gothic ever.

Other lexicographical work by Junius is contained in the edition of the *Gothic Gospels*, with glossary, he published with Thomas Marshall (1664–1665). His *Etymologicum Anglicanum* was only edited by Edward Lye in 1743. Projected books, partly contained in the more than 120 unpublished manuscripts he bequeathed to the Bodleian Library, Oxford (see article by E. G. Stanley in Bremmer 1998: 159–176), are glossaries on Chaucer, Otfrid, Cædmon, and Ælfric. There are also extant materials and lexicographical transcripts for a Latin–Old English dictionary (MSS Junius 113–116d) and his Old High German glossaries A-P (MSS Junius 116a-e). Later researchers draw heavily on his material (e.g., G. Hickes, *Linguarum veterum septentrionalium thesaurus* [1705], and E. Lye, *Dictionarium Saxonico-et Gothico-Latinum*, ed. O. Manning [2 vols., 1772]). For literature on Junius see Breuker (2000) and esp. Romburgh (2004).

His studies of Old and Middle Frisian, Anglo-Saxon, the Scandinavian languages, Gothic, Frankish, and Old High German earned him an important place in the early development of Old Germanic studies and comparative philology.

Bibliography

Bremmer R H Jr (ed.) (1998). *Franciscus Junius F. F. and his circle.* Amsterdam: Rodopi.

Breuker P H (2000). 'Junius, Franciscus.' In Hoops J (ed.) *Reallexikon der Germanischen Altertumskunde 16,* 2nd edn. Berlin: Mouton de Gruyter. 129–132.

Hickes G (1703–1705). *Linguarum vett. septentrionalium thesaurus grammatico-criticus et archaeologicus* (2 vols). London: Sheldon. Repr. Menston: Scolar Press 1970.

Junius F (1665). *Gothicum Glossarium, Quo pleraque Argentei Codicis vocabula explicantur, atque ex linguis cognatis illustrantur: Praemittuntur ei Gothicum, Runicum, Anglo-Saxonicum, aliáque Alphabeta.* Dordrecht: Junianus.

Junius F (1743). *Etymologicum Anglicanum.* Lye E (ed.). Oxford: Sheldon.

Junius F (1922). *Observations in Willerami Abbatis Francicam paraphrasin Cantici canticorum.* Voorwinden N (ed.). Amsterdam: Rodopi.

Junius F (2000). *Cædmonis monachi paraphrasis poetica: Genesios ac praecipuarum Sacrae paginae historiarum, abhinc annos M.LXX, Anglo-Saxonicè conscripta, & nunc primum edita.* Lucas P J (ed.). Amsterdam: Rodopi.

Junius F & Marshall Th (1655). *Quatuor D.N. Jesu Christi Euangeliorum versiones perantiquae duae, Gothica scil. et Anglo-Saxonica.* Dordrecht: Junianus, Essaeus.

Lye E (1772). *Dictionarium Saxonico et Gothico-Latinum* (2 vols). Manning O (ed.). London: Allen.

Kerling J (1984). 'Franciscus Junius, 17th-century lexicography and Middle English.' In Hartmann R R K (ed.) *LEXeter '83 proceedings.* Tübingen: Niemeyer. 92–100.

Mayou M B (1999). 'The sources of the Etymologicum Anglicanum (1743) by Franciscus Junius.' *Dictionaries 20,* 90–150.

Romburgh S van (2004). *For my worthy freind* [sic] *Mr Franciscus Junius: an edition of the correspondence of Francis Junius F. F. (1591–1677).* Leiden: Brill.

Jury Instructions

B Dumas, The University of Tennessee, Knoxville, TN, USA

Although jury trials are not used in all countries, they are used in most ADVERSARIAL or COMMON LAW jurisdictions. Those include Britain (apart from Scotland), continental Europe (including East Europe), the United States, Canada, Australia and New Zealand, as well as many other countries influenced by colonialism by those countries and continents. Trial process in all these jurisdictions includes the giving of instructions by the judge to the jury. In the United States, the instructions are almost always drawn from standard or pattern instructions, previously drafted models produced by bodies of lawyers and judges whose charge is to accurately state the current law. That is not often true in other jurisdictions. Further, the bulk of the research on the comprehensibility of jury instructions has taken place inside the United States. In many countries, jury proceedings are kept secret. Jurors are not allowed to discuss the deliberation process after the trial is concluded. For all these reasons, this article deals almost exclusively with the language of jury instructions used inside the United States. For information about other jury systems, readers may want to consult Vidmar (2000) and such articles as Thaman (1999).

Juries are used in both civil and criminal trials. Legal language in general has been recognized as being difficult for laypersons at least since Mellinkoff (1963) (see also Tiersma, 1999 and updates at his web page). But jury instructions present a special difficulty for jurors. In every jury trial, the trial judge, the expert on law, must instruct jurors, putative experts on fact, on the operative law. The judge must also instruct the jurors in how to apply the law to the evidence in the case. Jurors then must decide on blame or guilt, and the degree thereof, based on the evidence they have heard and the instructions they have received.

It has long been known that jurors do not always understand their instructions, particularly in the United States, where a great deal of research has been carried out by both judicial theorists and social scientists. Jurors may misunderstand the law or may misunderstand how they are to apply the law. Information about juror confusion has come from a number of sources. Case law (opinions) sometimes report the contents of notes sent by jurors to judges during deliberation; case law in some states allows testimony about conversations among jurors during deliberations; and of course, empirical evidence shows clearly that rewritten instructions that provide context, offer synonyms for difficult terms, and shorten sentences and modify grammatical constructions are much better understood than are pattern instructions.

Some lawyers and judges and also some social scientists assumed for a long time that jurors had difficulty understanding their instructions primarily because of the legal terms often used – often either Latin terms or terms of art, legal terms with specialized senses in law. When serious empirical study of juror comprehension by social scientists began (Charrow and Charrow, 1979), however, the findings indicated clearly that grammatical structures, including the structure of complex sentences, were a major source of problems. Although lay jurors are sometimes confused by Latin terms, such as *mens rea* and *res ipsa loquitur*, or legal terms of art, such as *infant* or *foreign corporation*, they are more often confused by specific grammatical structures and sentence complexity. In fact, the exact syntactic and semantic bars to jurors' comprehension of instructions are now well documented by linguists and psycholinguists. Researchers have demonstrated that modifying grammatical structures, simplifying sentence structure and giving additional information about the meanings of terms, particularly abstract terms, assist jurors to comprehend their instructions.

Further, jurors may be uncertain about the meanings of some highly abstract legal terms. One of the most difficult concepts to explain is that of *reasonable doubt*. Although the level of proof required in civil cases is a *preponderance of the evidence*, in a criminal case, it is *proof beyond a reasonable doubt*. The former can be quantified, but the latter cannot (see Tiersma, 1995; Dumas, 2002).

Legal systems have been slow to modify the pattern instructions used in most U.S. jurisdictions, partly because of the importance of precedent in law and partly because of judicial reluctance to relinquish any control over the jurors. Jury instructions are often worded to reflect the language in a statute, particularly in criminal cases (Solan, 2003), or an appellate opinion rather than to communicate a concept effectively to lay jurors. The judicial response to juror questions that reveal uncertainty about the meaning of jury instructions has typically been to respond by simply rereading the relevant portion of the instructions. There is, of course, no evidence that mere repetition of instructions is helpful when jurors report confusion. Indeed, some studies suggest that the typical judicial response to juror questions about

instructions treats the lack of comprehension as though it were a matter of not having heard, rather than not having understood.

History of Jury Instructions

The pattern or uniform jury instructions used in most U.S. jurisdictions represent in most cases an attempt on the part of the drafters to state the law accurately. Accuracy requires constitutional and statutory adequacy. Accuracy does not require comprehensibility by laypersons, and until recently such comprehensibility was not a primary goal of the drafters of pattern instructions.

The pattern instruction movement became influential in the late 1930s, when the judges of the Superior Court of Los Angeles published *The book of approved jury instructions* (generally referred to as the BAJI instructions). Other jurisdictions (Washington, D.C., Florida, Chicago, Nebraska, Colorado, and Utah) developed pattern instructions over the next two decades. Then in the mid-1950s, the Illinois Judicial Conference recommended the drafting of accurate and impartial instructions in order to correct some perceived abuses, and the recommendation was carried out under the supervision of the Illinois Supreme Court. Following that, the highest courts of a number of states followed the lead of Illinois and prepared their own pattern instructions. At this time, almost all U.S. states have pattern instructions, though the extent to which their use is mandated varies (see Nieland, 1979; Kelley and Wendt, 2002).

Pattern instructions were a step forward with respect to consistency and economy of time and effort. As they came into existence, they served as readily available instructions so that individual instructions did not have to be prepared for every legal case. Their weakness is their form; they are written in dense, complex language and are often modeled on appellate opinion language, language written by judges primarily for other judges to read. In general they are designed to be read, not to be heard read aloud. The judicial appeal of mirroring appellate language is that it is unlikely to give rise to a reversal for error. Put simply, jury instructions full of legalese are unlikely to be reversed for error by appellate courts if they use the same language used by those courts.

Pattern instructions can sometimes be modified by agreement of the parties, often as a result of a *charge conference* (sometimes called a *charging conference*), during which the judge meets with counsel for all parties and outlines the instructions. Counsel may request changes, though the judge makes all final decisions.

In recent years, some jurisdictions have taken a close look at problems in jury systems, particularly that of comprehensibility of instructions, and have initiated reform programs designed to increase comprehensibility and generally improve the lot of lay jurors. Some major studies are those by the National Center for State Courts (Munsterman *et al.*, 1997), California (Kelso, 1996), the District of Columbia (Council for Court Excellence, District of Columbia Project, 1998), New York (New York Jury Project, 1994), and Arizona (Dann, 1993). Two jurisdictions, Arizona and California, have made sweeping changes in their jury systems, including revision of instructions. Another jurisdiction, Tennessee, has implemented a number of changes based on the recommendations of a Jury Reform Commission (*Report of the Tennessee Bar Association Commission on Jury Reform,* 1999) and is continuing to consider suggestions (Cohen, 2003).

Similar recommendations have been made in all such studies. Pattern jury instructions present semantic and syntactic difficulties for lay jurors, difficulties that can be alleviated by careful rewriting of instructions. In addition, the methods of delivery of instructions need revision.

Comprehension Problems

The first major empirical study of pattern jury instructions was that of Charrow and Charrow (1979). In their study of 14 standard civil jury instructions (in Maryland), they tested comprehensibility by asking subjects to paraphrase the instructions, tested for understanding of legal concepts, and tested revised instructions. They found that the following grammatical constructions posed serious difficulties for many jurors:

1. Nominalizations
2. *As to* prepositional phrases
3. Unusual positions of phrases (removed from the elements they modify)
4. Deletion of *wh-* constructions and complements
5. Negatives, especially multiple negatives
6. Passives, particularly when used in subordinate clauses
7. Sentence complexity arising from multilayered subordination and embedded noun clauses.

In addition, Charrow and Charrow found that problems were presented by technical vocabulary, especially for abstract concepts such as *reasonable doubt*, and the use of doublets and triplets. Sentence length *per se* did not pose a problem – i.e., the number of words per sentence was not a good indicator of complexity. Revised instructions prepared to eliminate the

problems listed were much more comprehensible to jurors. All their results have been verified in subsequent research by other researchers.

It has been suggested (see Peter Tiersma's professional Web page) that if the mock jurors used in the Charrows' research had actually watched a trial, heard argument, and deliberated among themselves, juror comprehension might have been increased and might even have compensated for inadequate instructions. To test these and other points, Elwork, Sales, and Alfini (1982) reported results of research in which they showed a videotape of a criminal trial to volunteer jurors who served in mock juries. The mock juries watched the videotape, after which they were given either the original jury instructions or rewritten jury instructions. After verdicts were reached, the volunteers were questioned about legal points raised in the instructions. Those who had received the original instructions gave correct answers 40% of the time on average. Those who had received the revised instructions gave correct answers 78% of the time.

Another major study was that of Steele and Thornburg (1988), who, citing Charrow and Charrow (1979), addressed issues of comprehensibility from a judicial point of view, observing that the needed changes in jury instructions will "need to be supported by changes in the law that w[ill] create incentives for lawyers to worry as much about comprehensibility as they do about technical correctness" (p. 108).

Steele and Thornburg also addressed issues of the method of delivery of instructions, suggesting that jurors should be given copies of their instructions to use during deliberations.

It has, of course, been pointed out numerous times (see summaries in Dumas, 2000 and Gibbons, 2002) that the text of jury instructions is not the only factor affecting comprehension. Dumas has summarized delivery issues, taking the position that if the legal system had taken early research on the syntax and semantics of jury instructions totally seriously, they might well have made the suggested changes, but that if nothing else had changed, little improvement in comprehension might have occurred, and judges and lawyers might have concluded that the revisions were relatively unimportant. Methods of delivery and timing of delivery must also be considered (p. 67).

Gibbons has pointed out (drawing on research reported by Lieberman and Sales, among others), "There is ample research to show that linguistic issues are a part, but not all of the problem. There may also be psychological issues in jurors' understanding and acceptance of instruction" (2003: 179).

A forthcoming book (Dumas, 2005) suggests that changes are needed in all the following areas:

- Introductions to instructions (outlines, roadmaps)
- Syntactic complexity
- Semantic complexity
- Abstract concepts
- Timing and delivery
- Methods of delivery of jury instructions
- Instruction in the use of jury instructions.

Additional studies were cited earlier. They continue today. Of great importance are recent changes in California, particularly the approval of revised civil instructions, discussed in detail following.

Reform Efforts

In a sense, the creation of pattern instructions can be said to have begun the reform effort. However, because the wording of pattern instructions is responsible for most of the comprehension problems with jury instructions today, it makes more sense to date reform efforts from 1979, when Charrow and Charrow published their study. Serious efforts have thus continued for 25 years. Given the clear cross-disciplinary consensus about the difficulties of dealing with long, complex sentences and such terms of art as *mitigation* and *reasonable doubt* that has emerged, it is difficult to understand why the legal system has been so slow to make needed improvements to the pattern instructions. It must be attributed to ongoing judicial reluctance to relinquish control over the jurors.

But serious and quite successful efforts at reform have been made in several jurisdictions. Arizona was one of the first jurisdictions to make important and influential changes in the jury system as a whole. Most of the Arizona Committee's 55 recommendations were adopted by the Arizona Supreme Court in 1995; at least one follow-up study has reported positive results. Both the methods and conclusions of the reform commissions that carried out some of the studies are quite similar. In most cases, lawyers, judges, and jurors (the lawyers representing both prosecutors and defense lawyers in criminal cases and both plaintiffs' lawyers and defense lawyers in civil cases) studied the existing jury system and scholarly evaluations of that system, then made specific recommendations for change. Each recommended changes in many facets of jury procedures, though none has recommended the kinds of changes that will alter the traditional character of the American jury system.

The history of the Tennessee Bar Association is representative. It was created as a result of a presentation at an Annual Dinner of the National Center for State Courts at the Supreme Court Building in

Washington, D.C., in late 1997. TBA president-elect Pam Reeves heard Judge Michael Dann of Arizona describe the Arizona effort and similar jury reform projects in other states. After consulting with members of the Tennessee Supreme Court, Ms Reeves appointed the Tennessee Jury Reform Commission to review existing Tennessee jury practices, recent social science research, and developments in other jurisdictions and to make appropriate recommendations.

Each of three committees (jury administration and management, pretrial management and postverdict issues; jury selection and trial procedures; jury instructions) made recommendations, which were submitted to the entire Commission for discussion and approval. The conclusions of the Commission are contained in the May 1999 *Report of the Tennessee Bar Association Commission on Jury Reform* (hereafter *Report*). (See Cohen and Cohen, 2003 for a recent review of enactment of the recommendations.)

Reforms in New York and the District of Columbia have also been important. However, ongoing reform efforts in California deserve special attention. In July of 2003, the California Judicial Council unanimously approved approximately 800 new civil jury instructions and special verdict forms for use in California trial courts. The texts of the new instructions and verdict forms had been in preparation since 1997, when the Judicial Council created the Task Force on Jury Instructions in response to the 1996 statement by the Blue Ribbon Commission on Jury System Improvement to the effect that jury instructions would be more useful if they were stated in more understandable language. Though the new instructions are not required, their use is encouraged. (Details of the revision and comparisons of the new instructions with the older BAJI instructions are available at the website of the California Judicial Council. Some details are also available at Peter Tiersma's website.)

The Judicial Council also approved the preparation of new criminal instructions; they are expected to be submitted for approval by 2006.

Conclusion

Pattern instructions are not well understood by the average juror. The complexity of pattern instructions is multifaceted. Instructions are quite complex syntactically, and they should be rewritten so that they are less complex syntactically so as to be comprehensible to nonlawyers with limited formal education. But although it is true that revision of problematic linguistic constructions improves comprehensibility, that alone does not appear to be completely adequate.

In addition, there is evidence that a higher level of comprehension can be achieved by (1) providing roadmaps early in the instructions, (2) avoiding the use of arcane vocabulary items whenever possible, and (3) clarifying the meanings of difficult abstract concepts by providing concrete, perhaps narrative, examples.

All these steps are necessary both because of the inherent complexity of pattern instructions and perhaps also because of the nature of the Anglo-American legal process, which operates by means of a reasoning pattern in which a case is tried, a proposition descriptive of the case is made into a rule of law, and the proposition becomes a precedent to be applied to future similar situations. Future situations arise, some of which may eventually be sufficiently different from the original situation to prompt a rethinking of the original rule and perhaps a revision of it. It is thus essential that jurists be able to discern similarities and differences among objects and situations.

Roadmaps

During their time in the courtroom, jurors resemble nothing so much as students taking a course – or several courses simultaneously – in a new and arcane subject. They must absorb new information, learn new procedures, and digest and use new standards quickly and with no time for reading, reviewing, or consulting with others. They may, for example, need to learn a great deal about constitutional law, criminal procedure, medical standards, and human psychology all within a few days. In addition, they must judge the credibility of witnesses, often expert, who contradict each other's testimony. They need all the help they can get. One way to assist them in the instructions is to tell them exactly what they are going to hear before they hear all the details. A roadmap, a list or diagram telling jurors where they are headed and by what route that can expect to arrive there, can provide an outline of jury instructions.

Syntactic and Semantic Complexity

As noted earlier, pattern instructions often closely track statutory or appellate opinion language. Appellate opinions are specifically directed to judges (i.e., legal professionals trained in law and accustomed to using legal discourse on a daily basis). Further, they are designed to be read by judges, not listened to. They are necessarily characterized by lengthy sentences containing many clauses, a great deal of subordination, and arcane vocabulary items, including legal terminology and mysterious doublets. All those characteristics present comprehension problems.

Difficult Abstract Concepts

Many legal concepts are difficult to define; *reasonable doubt* is one such concept. Others are

proximate cause, defective and unreasonably dangerous, and *present cash value*. One way of explaining such concepts to lay jurors is to give examples or brief narratives. One possible corrective is paraphrasing. Another corrective involves the addition of an example. Some issues may require the addition of both positive and negative examples. Finally, some issues may suggest a need for both paraphrase and examples. Here is such an example making use of some of these strategies:

Some of these legal concepts or principles can be difficult for laypersons to understand. I hope this example will illustrate for you a practical example of proximate cause: it is negligence for a driver to drive a car that has bad, or slick, tires. If that automobile with bad tires slams into the rear of a car because the driver could not stop due to a combination of slick tires and wet pavement, then the negligence of the driver in driving with bad tires would be a proximate cause of the accident.

But it is possible for a person to be negligent without that negligence being a proximate cause of the accident. If the driver of that car with bad tires is stopped for a red light and is struck in the rear by another car, obviously the bad tires had nothing to do with the accident. In other words, the driver's negligence in driving a car with bad tires was not a *proximate cause* of that accident.

To find a party to be 'at fault,' you must find that party was negligent *and* that the negligence was a proximate cause of the injury or damage for which a claim was made. You must then determine the percentage of fault of each party whom you have determined is at fault.

Timing of the Delivery of the Instructions

It is traditional in most jurisdictions to provide most jury instructions, particularly those involving substantive law, after all the evidence has been presented and just before deliberation. This may invite jurors to form opinions about the proper outcome of the case well before they are instructed in the applicable law. It is widely believed that the single most important factor in the determination of juror attitude is juror reaction to the opening statements delivered at the very beginning of the case, when lawyers outline their theories of the case and prepare jurors for the evidence they will hear. Thus pretrial instruction in substantive law might improve jury performance.

Delivery of the Instructions

Whatever the exact manner of delivery, instructions prepared to be read should not be delivered orally in a rapid monotone. Jury instructions are often lengthy and complex. A number of changes in the delivery process could help to make the communication more effective. Judges could provide summaries as well as roadmaps (discussed earlier); they could make use of multimedia formats, particularly charts, decision trees, or simple 'bullet' summaries; and they could be careful to maintain eye contact with jurors and in general pay full attention to the legitimacy, dignity, and humanity of the jurors and their role.

Current research tells us that legal jargon is neither the sole nor a major source of incomprehensibility in jury instructions, though it can present special problems (*proximate cause*, *stipulate*, *mens rea*, *beyond a reasonable doubt*, etc.). The textual problem is overwhelmingly sentence structure, also difficult grammatical constructions such as nominalizations, vague or unusual prepositional phrases, doublets, especially of legal terms, and passive constructions, particularly in subordinate clauses. Other issues include presuppositions underlying jury instructions in opposition to the presuppositions underlying attorney strategies, the nature of the task, and the highly abstract nature of the instructions. Remedies are available: simplify sentence structure, paraphrase difficult concepts, add examples and narratives to the pattern instructions, answer jurors' questions, and give jurors written copies of their instructions.

Reform must acknowledge frankly that the jury's role as finder of fact has to compete both with requirements of legal procedure, particularly those implicating constitutional issues, and also with the judicial perception that the jury must be highly constrained and controlled. In the courtroom, the judge is the domain expert on law and procedure. Lip service is paid to the idea that the jury is the trier of fact and is, therefore, a domain expert in its own right. But jurors are not treated as experts. If they were, they would have greater freedom in certain areas, would be able to take notes and ask questions, and would be provided instructions that are not so arcane and convoluted as to be incomprehensible to most people.

See also: Legal Semiotics; Legal-Professional Language in Jury Trial.

Bibliography

Abbott W F & Batt J (eds.) (1999). *A handbook of jury research*. Philadelphia: American Law Institute-American Bar Association Committee on Continuing Professional Education.

Charrow R P & Charrow V R (1979). 'Making legal language understandable: a psycholinguistic study of jury instructions.' *Columbia Law Review 79*, 1306–1374.

Cohen N P & Cohen D R (2003). 'Jury reform in Tennessee.' *University of Memphis Law Review 34*, 1–71.

Council for Court Excellence, District of Columbia Project (1998). *Juries for the year 2000 and beyond: proposals to improve the jury systems in Washington, D.C.*
Dann M (1993). '"Learning lessons" and "speaking rights:" creating educated and democratic juries.' *Indiana Law Journal 68,* 1229–1277.
Dumas B K (2000a). 'US pattern jury instructions: problems and proposals.' *Forensic Linguistics: The International Journal of Language and the Law 7,* 49–71.
Dumas B K (2000b). 'Jury trials: lay jurors, pattern jury instructions, and comprehension issues.' *Tennessee Law Review 67,* 701–742.
Dumas B K (2002). 'Reasonable doubt about reasonable doubt: assessing jury instruction adequacy in a capital case.' In Cotterill J (ed.) *Language in the legal process.* Basingstoke and New York: Palgrave Macmillan Press. 246–259.
Dumas B K (Forthcoming 2005). *Writing and using effective jury instructions.* Notre Dame, Indiana: National Institute for Trial Advocacy.
Elwork A, Sales B D & Alfini J J (1982). *Making jury instructions understandable.* Charlottesville, VA: Michie.
Gibbons J (2003). 'Communication issues in the legal system.' Chapter 5 of *Forensic linguistics: an introduction to language in the justice system.* Malden, MA, and Oxford: Blackwell Publishing.
Garvey S P, Johnson S L & Marcus P (2000). 'Correcting deadly confusion: responding to jury inquiries in capital cases.' *Cornell Law Review 85,* 627–655.
Kassin S M & Wrightsman L S (1988). *The American jury on trial: psychological perspectives.* Oxford: Taylor and Francis.
Kelley P J & Wendt L A (2002). 'What judges tell juries about negligence: a review of pattern jury instructions.' *Chicago-Kent Law Review 77,* 587–623.
Kelso J C (1996). 'Final report of the Blue Ribbon Commission on Jury System Improvement.' *Hastings Law Journal 47,* 1433–1529.
Lieberman J & Sales B D (1999). 'The effectiveness of jury instructions.' In Abbott & Batt (eds.). 18-1–18-73.
Mellinkoff D (1963). *The language of the law.* Boston: Little, Brown.
Munsterman G T, Hannaford P L & Whitehead G M (eds.) (1997). *Jury trial innovations.* Williamsburg, VA: National Center for State Courts.
New York Jury Project (1994). Report to the Chief Judge of the State of New York, 31 March.
Nieland R G (1979). *Pattern jury instructions: a critical look at a modern movement to improve the jury system.* Chicago: American Judicature Society.
Solan L M (1999). 'Refocusing the burden of proof in criminal cases: some doubt about reasonable doubt.' *Texas Law Review 78,* 105–147.
Solan L M (2003). 'Jurors as statutory interpreters.' *Chicago-Kent Law Review 78,* 1281–1318.
Steele W W Jr & Thornburg E G (1988). 'Jury instructions: a persistent failure to communicate.' *North Carolina Law Review 67,* 77–119.
Stygall G (1994). *Trial language: differential discourse processing and discursive formation.* Amsterdam and Philadelphia: John Benjamins Publishing Company.
Tanford J A (1990). 'The law and psychology of Jury instructions.' *Nebraska Law Review 69,* 71–111.
Tanford J A (1991). 'Law reform by courts, legislatures, and commissions following empirical research on jury instructions.' *Law and Society Review 25, 155–175.*
Tennessee Bar Association Jury Reform Commission (1999). *Report of the Tennessee Bar Association Commission on Jury Reform.* Nashville, TN: Tennessee Bar Association. [Available at the Tennessee Bar Association website.]
Thaman S C (1999). 'Europe's new jury systems: the cases of Spain and Russia.' *Law and Contemporary Problems 62,* 233–259.
Tiersma P M (1993). 'Reforming the language of jury instructions.' *Hofstra Law Review 22,* 37–78.
Tiersma P M (1995). 'Dictionaries and death: do capital jurors understand mitigation?' *Utah Law Review 1,* 49.
Tiersma P (1999). *Legal language.* Chicago: University of Chicago Press.
Tiersma P (2001). 'The rocky road to legal reform: improving the language of jury instructions.' *Brooklyn Law Review 66,* 1081–1119.
Tiersma P M (2004). 'Jury instructions in the new millennium.' *Court Review 36,* 28–36.
Vidmar N (ed.) (2000). *World jury systems.* Oxford: Oxford University Press.

K

Kalkutungu

B J Blake, La Trobe University, Bundoora, Australia

Kalkutungu (also Kalkatungu (Kalkatung) and Kalkadoon) is a Pama-Nyungan language, and, like many other Australian languages, it is largely agglutinative, although with a good deal of irregular bound forms. Nouns are inflected for case and verbs for tense, aspect, mood, and modality. Word order is controlled by discourse principles with few grammatical constraints.

Kalkutungu has a typical phoneme inventory with two laminal series of stops, nasals, and laterals. In company with the Arandic languages to the west, at some stage in the past Kalkutungu underwent a change involving the loss of initial consonants and in some instances initial syllables. This has resulted in words with initial *a*, such as *arnka* 'to be ill' (cf. Yalarnnga *yarnka*) and words with initial homorganic nasal-stop clusters such as *mpaya* 'you two' (cf. Yalarnnga *nhumpala*).

The core case marking opposes an ergative case for the agent of a transitive verb (A) and an unmarked absolutive (alternatively nominative) for an intransitive subject (S) and object (O). The ergative-absolutive distinction applies to free pronouns as well as to nouns, but there is a system of clitic pronouns that oppose a subject set covering S and A functions to an object set. An odd feature of Kalkutungu is that although these clitics are obligatory with auxiliaries and in the perfect and imperfective aspects, they are otherwise optional and used rather sparingly. The following examples illustrate the core case marking and clitic pronouns. Note in (2) that the ergative expresses the instrument used to perform an action as well as the agent.

(1) marapayi ingka-nha-na ntiya-piyangu tawun-kunha
woman.ABS go-PT-3PL.SUBJ hill-ABL town-ALL
'the women went from the hill to the town'

(2) marapayi-thu ngayi kanimayi-nha-ngi-na rupu-ngku
woman-ERG me.ABS tie-PT-1SING.OBJ-3PL.SUBJ rope-ERG
'the women tied me up with rope'

There is a derived two-place intransitive construction, the antipassive, in which the subject appears in the absolutive and the object is demoted to the dative. In independent clauses, this construction is used to signal reduced semantic transitivity. It is used, for instance, to indicate a generic object as in (3b) rather than a specific object as in (3a), and it is obligatory in the imperfective.

(3a) marapai-thu rumpa-mi ithirr matyamirla-thu
woman-ERG grind-FU seed.ABS grindstone-ERG
'the woman will grind the seed with the grindstone'

(3b) marapai rumpa-yi-mi ithirr-ku matyamirla-thu
woman.ABS grind-AP-FU seed-DAT grindstone-ERG
'the woman will grind seed with the grindstone'

In dependent clauses, the antipassive is used to signal co-reference between the agent of the dependent clause and the S or O of the governing clause. In (4), (5), and (6), we have a type of subordinate clause used to express purpose adjuncts and complements to directive verbs. These are characterized by a purpose auxiliary, which hosts clitic pronouns. In (4), the agent of the purpose clause is co-referent with the S of the governing clause so the purpose clause appears in the antipassive, as is evident from the marking on the verb and the dative marking for the patient. In (5), the agent of the subordinate clause is co-referent with the O of the governing clause, so the antipassive is used. In (6), however, the agent of the purpose clause is coreferent with the allative-marked *mangarnaan-kunha* 'to the doctor', so no antipassive is used.

(4) ngayi ingka-nha nha-wu thuwarr-ku
I.ABS go-PT this-DAT snake-DAT
lhaa ngulurrma-yi
PURP.1SING.SUBJ catch-AP
'I went to catch the snake'

(5) nga-thu tjaa pati-nha thuku-wu
I-ERG this.ABS tell-PT dog-DAT
kuntu a-yi ngayima-yi
not PURP-3SING.SUBJ chase-AP
'I told this one not to chase the dog'

(6) nhaa nga-thu unpiyi nga-tji nhawurr
this.ABS I-ERG take I-DAT child.ABS
mangarnaan-kunha a-yi puthurr-puni
doctor-ALL PURP-3SING.SU good-make
'I took my child to the doctor for him/her to cure her'

A feature that Kalkutungu has in common with a number of other Australian languages is insubordination, whereby a subordinate clause becomes independent. Sentence (7a) provides an example; compare it with (4), (5), and (6). Essentially, a verb such as *ingka* in (4) is redundant and over time such verbs have been dropped in some circumstances to yield a construction in which the erstwhile dependent clause is independent. Witness the ergative on the agent in (7a), compared to the absolutive in (4). This construction, in which clitic pronouns are obligatory, makes an interesting comparison with (7b), in which the future tense is used and no clitics.

(7a) nyin-ti a-ngi lha?
you-ERG PURP-1SING.OBJ hit
'are you going to hit me?'

(7b) nhakaakuwa nyin-ti ngayi lhami?
why you-ERG me.ABS hit-FU
'why are you going to hit me?'

Kalkutungu has two applicative constructions. In one, a dative can be promoted to object. If there is a patient object, this is retained, so there is a double object construction. In sentence of (8a), the beneficiary is expressed in the dative. In (8b), the beneficiary is now O and is expressed by an object clitic.

(8a) utjan nga-tji intji-ya
firewood I-DAT chop-IMP
'chop some wood for me'

(8b) utjan intji-tjami-ya-ngi
firewood chop-APPL-IMP-1SING.OBJ
'chop me some wood'

In the other applicative, an instrumental or locative can be promoted. If there is a patient object, it is demoted to the dative. The principal function of this construction is to provide for the anaphoric deletion of the instrumental or locative in purpose or other subordinate clauses. Example (9) is typical. The knife is understood to be the instrument in the purpose clause, but it is covert. The patient is demoted to the dative. Note this demotion is not in response to coreference between the agent of the purpose clause and S or O, as in (4) and (5). Note too that the same marker *-nti* functions as a causative and applicative.

(9) kankari iti-nti-ya ati-ntji
knife. ABS return-CAUS-IMP meat-DAT
lhaa pintji-nti
PURP-1SING. SUBJ cut-AP
'bring the knife so I can cut the meat with it'

The final example is another sample of the instrumental applicative in which the instrument *utjula* 'net' has been promoted to O, pushing the patient *wakarri* 'fish' into the dative. It serves to illustrate two common principles of discourse. First, it is common to represent a nominal with a pronoun or other generic expression early in the clause with a more specific noun later. In (10), *nhaa* 'this' anticipates *utjula* 'net'. Second, the focus is normally placed first in the clause, in this instance *yawun* 'big'.

(10) yawun nhaa nga-thu utjula
big.ABS this.ABS I-ERG net.ABS
wakarri-yi ngurlurr-manti
fish-DAT catch-AP
'I use a big net to catch fish with'

See also: Australia: Language Situation; Australian Languages; Case; Ergativity; Head/Dependent Marking.

Bibliography

Blake B J (1979). *A Kalkatungu grammar.* Canberra, Australia: Pacific Linguistics.

Kambaata *See:* Highland East Cushitic Languages.

Kamei, Takashi (b.1912)

T Maruyama, Nanzan University, Nagoya, Japan

Born in Tokyo in 1912, Takashi Kamei graduated from the department of Japanese language and literature of the University of Tokyo in 1935, thereafter spending most of his career teaching linguistics at Hitotsubashi University in Tokyo. He also conducted academic researches and gave lectures at Cambridge University, Freie Universität Berlin from 1950 to 1953, and at the University of Arizona in 1975. He was also invited to give lectures on linguistics at various universities in Japan, including the University of Tokyo, Keio University, and International Christian University.

Most of Kamei's works are unknown to the academic world outside of Japan, mainly because most of them are written in Japanese. Some exceptions are 'A Saussurian mystery' (1982) and others listed below, but to some western scholars they might seem very complicated or tantalizing. Even Japanese scholars find it rather difficult to grasp what he really means owing to his unique writing style.

For Kamei, the historical study of language should not exist to compose and support the theoretical basis of linguistics, nor should it merely be the search for examples suited to establish such theories. It must be the pursuit of the development of a given language in its entirety, and in fact has that nature as its destiny. Expressing this from a negative point of view, the history of language can be nothing other than the history of a certain concrete individual language. From the point of view of comparative linguistics, French is one form of Neolatin. However, French is no language other than French and therefore is referred to as an individual subject, i.e., French, and by the same token it must not be considered as Latin. Metaphorically speaking, the moment Latin ceased to exist, French sprang to life. The interpretation of this kind of linguistic history cannot be made only considering a given language as a single code system. The work of the linguistic historian is therefore to take the detailed results of researchers in the field of diachronic linguistics and bring them once more correctly within the envelope of history.

In his own way, Kamei was comparable to Hugo Schuchardt in the sense that he always sat on the opposition benches vis-à-vis the mainstream of rationalism of linguistics of the 20th century, just like Schuchardt was always against the Neogrammarian way of thinking in the 19th century.

Kamei was one of the chief editors of *The Sanseido encyclopaedia of linguistics* (6 vols), published from 1988 to 1996 in Japan. It contains descriptions of more than 3500 different languages of the world and about 1500 important terminologies of linguistics. He also made a huge contribution to the understanding of the *seidaku* concept that had confronted scholars of Japanese for such a long time.

See also: Japan: History of Linguistics; Japanese; Schuchardt, Hugo (1842–1927).

Bibliography

Kamei T (1953). 'Are Japanese Turu and Ito related to Korean Turumi and Sil, respectively?' *The Annuals of the Hitotsubashi Academy III(2).*

Kamei T (1954). *Chinese borrowings in Preliterate Japanese.* Tokyo: Yoshikawa Kobunkan.

Kamei T (1954). 'On the authenticity of On-readings in Sino-Japanese.' *The Annuals of the Hitotsubashi Academy V(1).*

Kamei T (1957). 'On some extra-phonologic parallels to neutralisation proper.' *Travaux de l'Institut de linguistique II.*

Kamei T (1969). 'Betrachtungen eines Philologen über die Tennoherrschaft – Zum 100-jährigen Jubilaum der Meiji.' *Hitotsubashi Journal of Social Studies 5(1).*

Kamei T (1970). 'How did the sparrow twitter in ancient Japanese?' *Memoirs of the Research Department of the Toyo Bunko 28.*

Kamei T (1978). 'Covering and covered forms of women's language in Japanese.' *Hitotsubashi Journal of Arts and Sciences 19.*

Kamei T (1982). 'A Saussurian mystery.' In *Logos Semantikos, Studia Linguistica in Honorem Eugenio Coseriu,* vol. I. Berlin: W. de Gruyter.

Maruyama T & Barron J P (1988). 'Irrationalism in the historical study of language.' Tokyo University of Linguistic Papers '88.

Kamilaroi *See:* Gamilaraaya.

Kannada

S N Sridhar

This article is reproduced from the previous edition, volume 4, pp. 1833–1834, © 1994, Elsevier Ltd.

Kannada (formerly referred to sometimes as 'Kanarese' or 'Canarese') is a South Dravidian language spoken by about 28 million people primarily in the state of Karnataka, South India, where it is the official language. It is also spoken in the neighboring states of Tamil Nadu, Andhra Pradesh, Kerala, and Maharashtra.

History

Kannada is written in a variety of the 'alpha-syllabic' Brahmi system. The earliest written record, the Halmidi inscription (ca. 450 A.D.), already shows the influence of Sanskrit. The history of the language is conventionally divided into three periods, Old Kannada (up to ca. the thirteenth century A.D.), Middle Kannada (up to ca. the nineteenth century), and Modern Kannada (the nineteenth century onward).

Literary and Grammatical Tradition

Kannada has a world-class literary tradition, but few works have been translated. It also has a rich grammatical tradition. The first extant literary work, Nagavarma's *Kavirajamarga* (ninth century), is a treatise on poetics and includes the earliest grammar of Kannada. Kesiraja's *Śabdamaṇidarpana* (1260) is the classic grammar of Old Kannada. The other major grammar is Bhattakalanka's *Karna:taka Śabda:nuśa:sana* (1604), written in Sanskrit. Modern descriptions include Kittel (1903), Spencer (1950), Bright (1958), Schiffman (1983), and Sridhar (1989). Srikanthaiah (1960) and Bhat (1978) include other important descriptions in Kannada.

Variation

Three major regional varieties are recognized: the dialect of the former and present capitals – the Mysore/Bangalore Kannada, the Dharwar (northern) Kannada, and the Mangalore (coastal) Kannada. The Brahmin variety of Mysore Bangalore forms the basis of the standard language. It retains aspirated consonants and consonant cluster combinations in Sanskrit-derived words (which occur with high frequency in this variety). There are other caste varieties, which differ in phonology, morphology, syntax, and the lexicon. Kannada is a diglossic language.

Kannada is interesting from the point of view of historical sociolinguistics because of the millennium-old controversy between the elitists and populists on the one hand and the purists and the pragmatists on the other concerning the proper literary style – a debate that continues to this day.

Structure

Kannada morphology is agglutinative, suffixing, and quite regular. Nouns are marked by cases, postpositions, number, and occasionally gender. Verbs indicate tense, aspect, and agreement in number, person, and gender. Negation is expressed as main verb and as a suffix. Word order is subject–object–verb. The verb normally (but not always) ends the sentence. Modifiers (e.g., adjectives, adverbs, and subordinate clauses) precede their heads; auxiliaries follow main verbs. Compounding is very common in nouns and verbs. Normally, only one finite verb occurs per sentence: the syntax relies heavily on participles, gerunds, and infinitives. A number of these points are illustrated in the following sentence:

ni:nu	*nenne*	*meccida*	*a:*	*mu:ru*	*puṭṭa*
you	yesterday	admired	those	three	little

makkaḷu	*a:ṭa*	*a:ḍi*
children	game	play – past participle

daṇidu	*waḷamaneyalli*	*nidde*
tire – past participle	inside room – locative	sleep

ho:gidda:re.
go – past participle – continuous – 3 (human) plural
'Those three little children whom you admired yesterday have gone to sleep after playing and getting tired.'

Importance of Language Contact

Kannada has successfully assimilated an enormous amount of Sanskrit, Prakrit, Perso–Arabic, New Indo–Aryan, and more recently, English elements. This influence is manifested in the rather substantial stock of loan words, the productive use of Sanskrit derivational morphology, code-mixing, and calquing, which have been the preferred strategies for modernization throughout its history. This openness to borrowing sharply distinguishes Kannada from its sister language, Tamil.

Kannada and Linguistic Theory

Evidence from Kannada has played an influential role in several areas of modern linguistic inquiry. These include the status of grammatical relations, especially

nonnominative subjects, in syntactic theory; processes of syntactic and morphological convergence; morphological levels in derivation; caste dialects; and syntactic and psycholinguistic models of bilingual code-switching, among others.

See also: Dravidian Languages; India: Language Situation; Tamil Linguistic Tradition.

Bibliography

Bhat D N S (1978). *Kannada Vakyagalu: Antarika racane mattu Artha Vyavasthe.* Mysore: Geeta Press.

Bhattakalanka (1604). *Karnataka S'abda: nus'a: sanam.* Bangalore: Raj Kamal Prakashana.

Bright W (1958). *An Outline of Colloquial Kannada.* Poona: Deccan College Postgraduate and Research Institute.

Kesiraja (1260). *S'abdamanidarpana* Madras: University of Madras.

Kittel F (1903). *A Grammar of the Kannada language in English, comprising the three dialects of the language (Ancient, Mediaeval and Modern).* Mangalore: Basel Mission Book and Tract Depository.

Nagavarma (ca. 9th century). *Karnataka Bhasha Bhushana.* New Delhi: Asia Educational Services.

Schiffman H F (1983). *A reference grammar of spoken Kannada.* Seattle: University of Washington Press.

Spencer H (1950). *A Kanarese grammar.* Mysore: Wesley Press.

Sridhar S N (1989). *Kannada.* London: Routledge.

Srikanthaiah T N (1960). *Kannada Madhyama Vya:karana.* Gubbi: Sevasadana Press.

Kant, Immanuel (1724–1804)

K Willems, University of Ghent, Ghent, Belgium

Kant was born in Königsberg (now called Kaliningrad and a part of Russia), at the time the administrative center of East Prussia. In 1740 he entered the university of his native city to study theology, philosophy, physics, and mathematics. After receiving his doctorate in 1755, he became a *Privatdozent* and earned his living mainly as a private tutor. He never left the area of Königsberg during his entire life. In 1770 he was finally appointed professor of logic and metaphysics at the university of Königsberg after refusing several offers from other universities since 1764. He held his position until 1796, serving as rector of the university in 1786 and 1788.

Kant has often been ranked among the most influential thinkers of all time, on a par with Plato, Aristotle, St. Thomas, Spinoza, Leibniz, Hegel, Nietzsche, and others. His 'transcendental philosophy' has had a long-lasting impact on the history of philosophy and of Western thought in general. It is considered a groundbreaking epistemological framework for assessing the aims, methods, and limits of human knowledge, including moral and aesthetic judgments. Kant's fame is based mainly on his three major works, his so-called *Critiques*, viz., the *Critique of pure reason* (1781, second revised edition 1787), the *Critique of practical reason* (1788), and the *Critique of judgment* (1790). Apart from these books, however, Kant wrote a great number of treatises on a variety of subjects, including natural history, theology, religion, and law.

Kant's philosophy marks the beginning of 'idealism' and is the single most influential synthesis of rationalist and Enlightenment thought on the one hand and empiricist considerations (i.e., a realistic stance) on the other. According to Kant, thinking (understanding) and perceiving are two distinct faculties of the mind, which are related in a complex way. He believed that everything human beings can know about the external world goes back to mental acts in which data are perceived or experienced according to rules and forms of understanding imposed on the data by the mind itself. For Kant, both the data and the rules and forms of understanding are real, but the mind must impose its own constraints on the data for experience to be possible. Although essential component parts of human knowledge are not derived from experience (on this issue Kant is in opposition to the philosophy of David Hume), they are ultimately inconceivable without at least presuming the possibility of their being applied to data in experience. For example, assessing a causal relation between two events in the world means, according to Kant, that causality is projected by the mind onto the events and cannot be found in the latter as such, but at the same time causality is nothing without real-world events to which it can be applied. Thus the world we experience is not simply a product of our minds; rather, our knowledge is confined to

nonnominative subjects, in syntactic theory; processes of syntactic and morphological convergence; morphological levels in derivation; caste dialects; and syntactic and psycholinguistic models of bilingual code-switching, among others.

See also: Dravidian Languages; India: Language Situation; Tamil Linguistic Tradition.

Bibliography

Bhat D N S (1978). *Kannada Vakyagalu: Antarika racane mattu Artha Vyavasthe*. Mysore: Geeta Press.

Bhattakalanka (1604). *Karnataka S'abda: nus'a: sanam*. Bangalore: Raj Kamal Prakashana.

Bright W (1958). *An Outline of Colloquial Kannada*. Poona: Deccan College Postgraduate and Research Institute.

Kesiraja (1260). *S'abdamanidarpana* Madras: University of Madras.

Kittel F (1903). *A Grammar of the Kannada language in English, comprising the three dialects of the language (Ancient, Mediaeval and Modern)*. Mangalore: Basel Mission Book and Tract Depository.

Nagavarma (ca. 9th century). *Karnataka Bhasha Bhushana*. New Delhi: Asia Educational Services.

Schiffman H F (1983). *A reference grammar of spoken Kannada*. Seattle: University of Washington Press.

Spencer H (1950). *A Kanarese grammar.* Mysore: Wesley Press.

Sridhar S N (1989). *Kannada*. London: Routledge.

Srikanthaiah T N (1960). *Kannada Madhyama Vya:karana*. Gubbi: Sevasadana Press.

Kant, Immanuel (1724–1804)

K Willems, University of Ghent, Ghent, Belgium

Kant was born in Königsberg (now called Kaliningrad and a part of Russia), at the time the administrative center of East Prussia. In 1740 he entered the university of his native city to study theology, philosophy, physics, and mathematics. After receiving his doctorate in 1755, he became a *Privatdozent* and earned his living mainly as a private tutor. He never left the area of Königsberg during his entire life. In 1770 he was finally appointed professor of logic and metaphysics at the university of Königsberg after refusing several offers from other universities since 1764. He held his position until 1796, serving as rector of the university in 1786 and 1788.

Kant has often been ranked among the most influential thinkers of all time, on a par with Plato, Aristotle, St. Thomas, Spinoza, Leibniz, Hegel, Nietzsche, and others. His ‘transcendental philosophy’ has had a long-lasting impact on the history of philosophy and of Western thought in general. It is considered a groundbreaking epistemological framework for assessing the aims, methods, and limits of human knowledge, including moral and aesthetic judgments. Kant’s fame is based mainly on his three major works, his so-called *Critiques*, viz., the *Critique of pure reason* (1781, second revised edition 1787), the *Critique of practical reason* (1788), and the *Critique of judgment* (1790). Apart from these books, however, Kant wrote a great number of treatises on a variety of subjects, including natural history, theology, religion, and law.

Kant’s philosophy marks the beginning of ‘idealism’ and is the single most influential synthesis of rationalist and Enlightenment thought on the one hand and empiricist considerations (i.e., a realistic stance) on the other. According to Kant, thinking (understanding) and perceiving are two distinct faculties of the mind, which are related in a complex way. He believed that everything human beings can know about the external world goes back to mental acts in which data are perceived or experienced according to rules and forms of understanding imposed on the data by the mind itself. For Kant, both the data and the rules and forms of understanding are real, but the mind must impose its own constraints on the data for experience to be possible. Although essential component parts of human knowledge are not derived from experience (on this issue Kant is in opposition to the philosophy of David Hume), they are ultimately inconceivable without at least presuming the possibility of their being applied to data in experience. For example, assessing a causal relation between two events in the world means, according to Kant, that causality is projected by the mind onto the events and cannot be found in the latter as such, but at the same time causality is nothing without real-world events to which it can be applied. Thus the world we experience is not simply a product of our minds; rather, our knowledge is confined to

what we can conceive in the first place (the 'phenomena'). On the other hand, we also have to presuppose a world in its own right that is wholly external to us and consists of things as they are in themselves (called 'noumena' or *Dinge an sich* by Kant), about which we can know nothing.

Although he presupposes semiotic functions for language throughout his work, Kant does not explicitly address language for its own sake in his writings, apart from a few passages on language in the *Critiques* and in sections 38–39 of his 1798 treatise on *Anthropology.* The fact that Kant paid so little attention to language in his work was criticized severely by some of his contemporaries, in particular his former student Johann Gottfried Herder (the author of one of the most famous essays on *The origin of language*, 1772). Interestingly, Kant never found it necessary to respond to this criticism. However, Kant's philosophy eventually proved important historically for both the philosophy of language and the epistemology of linguistics (*see* **Humboldt, Wilhelm von (1767–1835)**). Well known in this regard is Kant's new list of 'categories' (*see* **Aristotle and Linguistics**). These categories are defined as "pure concepts of understanding." According to Kant, they are indispensable to experience and common sense as well as to scientific knowledge. Kant proposes a list of 12 categories, which he subdivides into four classes (quantity, quality, relation, and modality); they include basic concepts of reasoning, such as unity, totality, negation, causality, possibility, necessity, etc. Like the Aristotelian categorical concepts, the Kantian categories have a strong bearing on linguistic forms: not only do they correspond to concepts inevitably couched in words and phrases in Kant's philosophical system; they also show that language and reason mutually presuppose each other. (Observe that the four Gricean conversational maxims of quantity, quality, relation, and manner refer, albeit in a loose sense, to Kant's system of categories. (*see* **Grice, Herbert Paul (1913–1988)**.) Finally, Kant's famous logical distinction between 'analytic' and 'synthetic' judgments is also intimately linked to the linguistic form of propositions. Whereas analytic propositions add nothing to the concepts one already has (and therefore cannot be denied), synthetic propositions can be achieved only by appealing to something outside them, i.e., experience or intuition, thus supplying new information, which is in principle falsifiable (*see* **Quine, Willard van Orman (1908–2000)**).

See also: Aristotle and Linguistics; Grice, Herbert Paul (1913–1988); Humboldt, Wilhelm von (1767–1835); Quine, Willard van Orman (1908–2000); Thought and Language: Philosophical Aspects.

Bibliography

Guyer P (ed.) (1992). *The Cambridge companion to Kant.* Cambridge: Cambridge University Press.

Höffe O (1994). *Immanuel Kant.* New York: State University of New York Press.

Kant I (1781). *Kritik der reinen Vernunft* (1st edn.). [2nd edn., 1787; English translation: *Critique of pure reason.* New York: St. Martin's Press (1965) and Cambridge: Cambridge University Press (1999).] Riga: J F Hartknoch.

Kant I (1783). *Prolegomena zu einer jeden künftigen Metaphysik, die als Wissenschaft wird auftreten können.* [English translation: *Prolegomena to any future metaphysics.* Indianapolis: Hackett (1977) and Cambridge: Cambridge University Press (1997).] Riga: J F Hartknoch.

Kant I (1788). *Kritik der praktischen Vernunft.* [English translation: *Critique of practical reason.* New York: Garland (1976).] Riga: J F Hartknoch.

Kant I (1790). *Kritik der Urteilskraft.* [English translation: *Critique of judgment.* Indianapolis: Hackett (1987).] Berlin and Liebau: Lagarde & Friederich.

Kant I (1798). *Anthropologie in pragmatischer Hinsicht.* [English translation: *Anthropology from a pragmatic point of view.* The Hague: Martinus Nijhoff (1974).] Königsberg: F Nicolovius.

Kant-Studien. Philosophische Zeitschrift der Kant-Gesellschaft. Berlin/New York: Walter de Gruyter [published since 1896; four issues per year].

Kanuri

G J Dimmendaal, University of Cologne, Cologne, Germany

One of the earliest documents on an African language, a short vocabulary dating back to the 17th century, involves data from Kanuri, also known as Yerwa, or Bornu; the name 'Beriberi' is considered to be derogatory by Kanuri speakers. Kanuri is a major language in terms of the number of speakers; estimates for Nigeria range between three and four million, whereas in neighboring countries like Cameroon, Chad, and Niger, there may be around half a million speakers. Kanuri forms a dialect cluster with Kanembu, which is spoken by a distinct ethnic

what we can conceive in the first place (the 'phenomena'). On the other hand, we also have to presuppose a world in its own right that is wholly external to us and consists of things as they are in themselves (called 'noumena' or *Dinge an sich* by Kant), about which we can know nothing.

Although he presupposes semiotic functions for language throughout his work, Kant does not explicitly address language for its own sake in his writings, apart from a few passages on language in the *Critiques* and in sections 38–39 of his 1798 treatise on *Anthropology.* The fact that Kant paid so little attention to language in his work was criticized severely by some of his contemporaries, in particular his former student Johann Gottfried Herder (the author of one of the most famous essays on *The origin of language*, 1772). Interestingly, Kant never found it necessary to respond to this criticism. However, Kant's philosophy eventually proved important historically for both the philosophy of language and the epistemology of linguistics (*see* **Humboldt, Wilhelm von (1767–1835)**). Well known in this regard is Kant's new list of 'categories' (*see* **Aristotle and Linguistics**). These categories are defined as "pure concepts of understanding." According to Kant, they are indispensable to experience and common sense as well as to scientific knowledge. Kant proposes a list of 12 categories, which he subdivides into four classes (quantity, quality, relation, and modality); they include basic concepts of reasoning, such as unity, totality, negation, causality, possibility, necessity, etc. Like the Aristotelian categorical concepts, the Kantian categories have a strong bearing on linguistic forms: not only do they correspond to concepts inevitably couched in words and phrases in Kant's philosophical system; they also show that language and reason mutually presuppose each other. (Observe that the four Gricean conversational maxims of quantity, quality, relation, and manner refer, albeit in a loose sense, to Kant's system of categories. (*see* **Grice, Herbert Paul (1913–1988)**.) Finally, Kant's famous logical distinction between 'analytic' and 'synthetic' judgments is also intimately linked to the linguistic form of propositions. Whereas analytic propositions add nothing to the concepts one already has (and therefore cannot be denied), synthetic propositions can be achieved only by appealing to something outside them, i.e., experience or intuition, thus supplying new information, which is in principle falsifiable (*see* **Quine, Willard van Orman (1908–2000)**).

See also: Aristotle and Linguistics; Grice, Herbert Paul (1913–1988); Humboldt, Wilhelm von (1767–1835); Quine, Willard van Orman (1908–2000); Thought and Language: Philosophical Aspects.

Bibliography

Guyer P (ed.) (1992). *The Cambridge companion to Kant.* Cambridge: Cambridge University Press.

Höffe O (1994). *Immanuel Kant.* New York: State University of New York Press.

Kant I (1781). *Kritik der reinen Vernunft* (1st edn.). [2nd edn., 1787; English translation: *Critique of pure reason.* New York: St. Martin's Press (1965) and Cambridge: Cambridge University Press (1999).] Riga: J F Hartknoch.

Kant I (1783). *Prolegomena zu einer jeden künftigen Metaphysik, die als Wissenschaft wird auftreten können.* [English translation: *Prolegomena to any future metaphysics.* Indianapolis: Hackett (1977) and Cambridge: Cambridge University Press (1997).] Riga: J F Hartknoch.

Kant I (1788). *Kritik der praktischen Vernunft.* [English translation: *Critique of practical reason.* New York: Garland (1976).] Riga: J F Hartknoch.

Kant I (1790). *Kritik der Urteilskraft.* [English translation: *Critique of judgment.* Indianapolis: Hackett (1987).] Berlin and Liebau: Lagarde & Friederich.

Kant I (1798). *Anthropologie in pragmatischer Hinsicht.* [English translation: *Anthropology from a pragmatic point of view.* The Hague: Martinus Nijhoff (1974).] Königsberg: F Nicolovius.

Kant-Studien. Philosophische Zeitschrift der Kant-Gesellschaft. Berlin/New York: Walter de Gruyter [published since 1896; four issues per year].

Kanuri

G J Dimmendaal, University of Cologne, Cologne, Germany

One of the earliest documents on an African language, a short vocabulary dating back to the 17th century, involves data from Kanuri, also known as Yerwa, or Bornu; the name 'Beriberi' is considered to be derogatory by Kanuri speakers. Kanuri is a major language in terms of the number of speakers; estimates for Nigeria range between three and four million, whereas in neighboring countries like Cameroon, Chad, and Niger, there may be around half a million speakers. Kanuri forms a dialect cluster with Kanembu, which is spoken by a distinct ethnic

group mainly in Chad. Kanuri-Kanembu is part of the Saharan family, a well-defined subgroup within the Nilo-Saharan phylum.

Kanuri was intimately linked with the Kanem-Borno empire for almost 1000 years. Its role as the *lingua franca* of northern Nigeria was reduced in favor of Hausa during colonial times. Parallel to Hausa, a modified Arabic script known as Ajami was used in Kanuri for several centuries. The orthography based on Roman script and developed during the British colonial period was standardized in 1974. Today, Kanuri is used in mother-tongue education as well as on the radio and in television in Nigeria. It is also taught at university level, e.g., at the University of Maiduguri (Nigeria).

One of the earliest detailed analyses of Kanuri is Lukas (1937), who had already showed that Kanuri has two distinctive tone levels, low and high, as in *kànúrì* ('Kanuri person') or *kànùrí* ('Kanuri language'). These register tones may also be combined to build complex tones; the so-called 'mid' tone established by Lukas (1937: 3) represents a downstep high tone, i.e., a conditioned variant of the high tone. As further shown in Lukas' pioneering study, as well as in more recent analyses by Hutchison (1981), and by Cyffer and Geider (1997), Kanuri also has an intricate system of consonant alternation, resulting in complex morphophonemic alternations.

One interesting typological property of Kanuri is the relatively small number of verbs. More than 95% of the predicative structures are formed by a combination of the light verb *n* ('say, think') and some complement (for example: *le+n* ('go')). This latter strategy is common in a range of Nilo-Saharan languages stretching from Kanuri in the Lake Chad region to Nara in Eritrea (*see* **Nilo-Saharan Languages**). Both complex predicates of this type and basic verbs are inflected for subject, tense-aspect-mood, as well as object (in the case of first and second person). Derivational argument modulation (more specifically, neutro-passive, causative, applicative marking) and pluractional marking is also expressed in the verbal complex.

Kanuri is head-final at the clausal level, involving verb-final structures, preverbal complementation, and the use of postpositions, but nominal modifiers follow the head noun. From a historical-comparative point of view, Kanuri grammar appears to be characteristic for the Saharan group within Nilo-Saharan. Interestingly, however, the lexical structure of Saharan languages appears to be less stable, as argued by Cyffer (2000).

See also: Nigeria: Language Situation; Nilo-Saharan Languages.

Bibliography

Cyffer N (2000). 'Linguistic properties of the Saharan languages.' In Zima P (ed.) *Areal and genetic factors in language classification and description: Africa south of the Sahara.* Munich: Lincom Europa. 30–59.

Cyffer N & Geider Th (eds.) (1997). *Advances in Kanuri scholarship.* Cologne: Rüdiger Köppe.

Cyffer N & Hutchison J P (1990). *Dictionary of the Kanuri language.* Dordrecht and Maiduguri: Foris Publications and University of Maiduguri.

Hutchison J P (1981). *The Kanuri language: A reference grammar.* Madison: University of Wisconsin.

Lukas J (1937). *A study of the Kanuri language.* London: Oxford University Press.

Kapampangan

H Kitano, Aichi University of Education, Aichi, Japan

Kapampangan (Pampangan) is spoken mainly in Pampanga Province and in parts of Tarlac, Nueva Ecija, Bulacan, and Bataan provinces of Luzon, the Philippines. It is one of the largest languages in the Philippines. The number of its speakers is estimated as 1 897 378 (1990 census).

Phonology

Kapampangan has the consonants [p, b, t, d, k, g, ʔ, s, tʃ, ʤ, m, n, ŋ, l, r, w, y] and vowels [i, e, a, ø, u].

Overview of Kapampangan Grammar

Word Order

Like other Philippine languages, Kapampangan is a predicate-initial language. Clitic pronouns and clitic adverbs usually occupy the second position of a clause, and clitic pronouns are almost always obligatory, even when their coreferent noun phrases are present.

(1) Masakit-ya ing lalaki
sick-ABS.3SG DET.ABS.SG man
'the man is sick'

Case Marking

Kapampangan has three cases: topic, genitive (or nontopic), and oblique. Since Kapampangan exhibits

group mainly in Chad. Kanuri-Kanembu is part of the Saharan family, a well-defined subgroup within the Nilo-Saharan phylum.

Kanuri was intimately linked with the Kanem-Borno empire for almost 1000 years. Its role as the *lingua franca* of northern Nigeria was reduced in favor of Hausa during colonial times. Parallel to Hausa, a modified Arabic script known as Ajami was used in Kanuri for several centuries. The orthography based on Roman script and developed during the British colonial period was standardized in 1974. Today, Kanuri is used in mother-tongue education as well as on the radio and in television in Nigeria. It is also taught at university level, e.g., at the University of Maiduguri (Nigeria).

One of the earliest detailed analyses of Kanuri is Lukas (1937), who had already showed that Kanuri has two distinctive tone levels, low and high, as in *kànúrì* ('Kanuri person') or *kànùrí* ('Kanuri language'). These register tones may also be combined to build complex tones; the so-called 'mid' tone established by Lukas (1937: 3) represents a downstep high tone, i.e., a conditioned variant of the high tone. As further shown in Lukas' pioneering study, as well as in more recent analyses by Hutchison (1981), and by Cyffer and Geider (1997), Kanuri also has an intricate system of consonant alternation, resulting in complex morphophonemic alternations.

One interesting typological property of Kanuri is the relatively small number of verbs. More than 95% of the predicative structures are formed by a combination of the light verb *n* ('say, think') and some complement (for example: *le+n* ('go')). This latter strategy is common in a range of Nilo-Saharan languages stretching from Kanuri in the Lake Chad region to Nara in Eritrea (*see* **Nilo-Saharan Languages**). Both complex predicates of this type and basic verbs are inflected for subject, tense-aspect-mood, as well as object (in the case of first and second person). Derivational argument modulation (more specifically, neutro-passive, causative, applicative marking) and pluractional marking is also expressed in the verbal complex.

Kanuri is head-final at the clausal level, involving verb-final structures, preverbal complementation, and the use of postpositions, but nominal modifiers follow the head noun. From a historical-comparative point of view, Kanuri grammar appears to be characteristic for the Saharan group within Nilo-Saharan. Interestingly, however, the lexical structure of Saharan languages appears to be less stable, as argued by Cyffer (2000).

See also: Nigeria: Language Situation; Nilo-Saharan Languages.

Bibliography

Cyffer N (2000). 'Linguistic properties of the Saharan languages.' In Zima P (ed.) *Areal and genetic factors in language classification and description: Africa south of the Sahara.* Munich: Lincom Europa. 30–59.

Cyffer N & Geider Th (eds.) (1997). *Advances in Kanuri scholarship.* Cologne: Rüdiger Köppe.

Cyffer N & Hutchison J P (1990). *Dictionary of the Kanuri language.* Dordrecht and Maiduguri: Foris Publications and University of Maiduguri.

Hutchison J P (1981). *The Kanuri language: A reference grammar.* Madison: University of Wisconsin.

Lukas J (1937). *A study of the Kanuri language.* London: Oxford University Press.

Kapampangan

H Kitano, Aichi University of Education, Aichi, Japan

Kapampangan (Pampangan) is spoken mainly in Pampanga Province and in parts of Tarlac, Nueva Ecija, Bulacan, and Bataan provinces of Luzon, the Philippines. It is one of the largest languages in the Philippines. The number of its speakers is estimated as 1 897 378 (1990 census).

Phonology

Kapampangan has the consonants [p, b, t, d, k, g, ʔ, s, tʃ, ʤ, m, n, ŋ, l, r, w, y] and vowels [i, e, a, ø, u].

Overview of Kapampangan Grammar

Word Order

Like other Philippine languages, Kapampangan is a predicate-initial language. Clitic pronouns and clitic adverbs usually occupy the second position of a clause, and clitic pronouns are almost always obligatory, even when their coreferent noun phrases are present.

(1) Masakit-ya ing lalaki
sick-ABS.3SG DET.ABS.SG man
'the man is sick'

Case Marking

Kapampangan has three cases: topic, genitive (or nontopic), and oblique. Since Kapampangan exhibits

an ergative system in pronominal case marking, these cases may also be called absolutive, ergative, and oblique, respectively, which are the terms used in this article.

Negation

Sentential negation is expressed by the predicate-initial *e*.

(2) E-ya masakit ing lalaki
NEG-ABS.3SG sick DET.ABS.SG man
'The man is not sick'

Existence and Possession

Both existence and possession are expressed by the existential particles *atin(g)* ('there is, be present, have') and *ala* ('there is not, be absent, do not have'). (LK = linker, realized either as = *ng* [after vowels, *n*, and the glottal stop] or *a* elsewhere)

(3) Ating metung a babai-ng malagu
EXIST one LK woman-LK beautiful
'there was a beautiful woman'

(4) Atin-ya-ng kapatad a lalaki
EXIST-ABS.3SG-LK sibling LK man
'he/she has a brother (male sibling)'

Lexical Classes

Nouns Nouns may be monomorphemic or derived by affixation. A limited number of nouns may be pluralized with vowel lengthening (e.g., *lalá:ki* 'man' and *la:lá:ki* 'men,' *babá:i* 'woman' and *ba:bá:i* 'women').

Noun phrases are formed with other lexical classes, such as adjectives and verbs, by use of the linker.

(5) ing bayu-mu-ng imalan
DET.ABS.SG new-ERG.2SG-LK dress
'your new dress'

(6) ing balita-ku-ng dimdam kang Mike
DET.ABS.SG news-ERG.1SG-LK heard DET.OBL.SG Mike
'the news I heard from Mike'

Table 1 Determiners in Kapampangan

	Absolutive	*Ergative*	*Oblique*
Common nouns			
SG	ing, = ng	ning	king, keng
PL	deng/reng	reng	karing
Personal names			
SG	i	= ng	kang
PL	di/ri	ri	kari

Verbs Verbs derive for focus (actor, patient, directional, beneficiary, and instrumental; see section on focus constructions) and inflect for aspect (contingent [contemplated, future], perfective [completed, past], and imperfective [incompleted, progressive, present]). Verbs derive by way of affixation (prefixes, suffixes, and infixes), reduplication, vowel alternations, vowel lengthening, and combinations thereof. Some other affixes denote causative, aptative (abilitative, accidental, and coincidental actions), and distributive (states or actions distributed over space and time, and repetitive actions).

Adjectives Adjectives may be monomorphemic or derived, with affixes, such as *ma-* (plural *manga-*), which is the most common adjective-forming prefix. The comparative degree is marked by *mas* (e.g., *mas maragul* 'bigger'). Superlative adjectives are formed with the prefix *peka-* (*pekamaragul* 'biggest').

Determiners Determiners, also called articles, case markers, or noun markers, prenominally indicate the case (absolutive, ergative, or oblique) and number (singular or plural) of the nouns they qualify, and whether the nouns are common or personal (Table 1).

Pronouns Pronouns may be clitic or free. Some combinations of two clitic pronouns (ergative and absolutive) may be fused into one word; e.g., *na* (ERG.1SG) + *ya* (ABS.3SG) becomes *ne* (Table 2).

Focus Constructions

In Kapampangan, as in other Philippine languages, the morphology of the verb indicates the semantic relationship between the predicate and the absolutive argument. In an actor-focus construction, the absolutive argument is semantically an actor, and the verbal predicate takes appropriate actor-focus affixes. Likewise, in a patient-focus construction, the absolutive argument is semantically a patient, and the verbal predicate takes patient-focus affixes. In the following examples of each focus construction, the boldfaced argument is the absolutive.

Actor focus:

(7) Mamangan la-ng manuk
eating ABS.3 PL-LK chicken
'**they** are eating chicken'

Patient focus:

(8) Kakanan de ing manuk
eating ERG.3PL + ABS.3SG DET.ABS.SG chicken
'they are eating **the chicken**'

Table 2 Pronouns in Kapampangan

	ABS (clitic)	ABS (free)	ERG (clitic)	OBL (free)
1 SG	ku	yaku, aku	ku	kanaku, kaku
2 SG	ka	ika	mu	keka
3 SG	ya	iya	na	kaya
1 DUAL	kata	ikata	ta	kekata
1 PL INCL	katamu, tamu, kata, ta	ikatamu, itamu, ikata, ita	tamu, ta	kekatamu, kekata
1 PL EXCL	kami, ke	ikami, ike	mi	kekami, keke
2 PL	kayu, ko	ikayu, iko	yu	kekayu, keko
3 PL	la	ila	da/ra	karela

Directional focus:

(9) Dinan me-ng
give.to ERG.2SG + ABS.3SG-LK
pera ita-ng anak
money that.ABS.SG-LK child
'give some money **to that kid**'

Beneficiary focus:

(10) Pangadi me
pray.for ERG.2SG + ABS.3SG
'pray **for him/her/it**'

Instrumental focus:

(11) Penyulat ne
wrote.with ERG.3SG + ABS.3SG
ini-ng lapis
this.ABS.SG-LK pencil
'he/she wrote **with this pencil**'

See also: Clitics; Ergativity; Malayo-Polynesian Languages; North Philippine Languages; Philippines: Language Situation; Pluractionals (Distributives); Reduplication; Southeast Asia as a Linguistic Area; Tagalog.

Bibliography

Forman M L (1971a). *Kapampangan grammar notes.* Honolulu: University of Hawaii Press.

Forman M L (1971b). *Kapampangan dictionary.* Honolulu: University of Hawaii Press.

Gonzalez A B (1981). *Pacific linguistics, series C, no. 48: Pampangan: towards a meaning-based description.* Canberra: Department of Linguistics, The Australian National University.

Grimes B F (ed.) (2000). *Ethnologue* (14th edn.). Available at www.ethnologue.com.

Kitano H (2001). 'Kapampangan.' In Garry J & Rubino C (eds.) *Facts about the world's languages: an encyclopedia of the world's major languages, past and present.* New York/Dublin: H. W. Wilson Company. 371–375.

Mirikitani L T (1971). *Speaking Kapampangan.* Honolulu: University of Hawaii Press.

Mirikitani L T (1972). *Kapampangan syntax.* Honolulu: University of Hawaii Press.

Mithun M (1994). 'The implications of ergativity for a Philippine voice system.' In Fox B & Hopper P J (eds.) *Voice: form and function.* Amsterdam: John Benjamins. 247–277.

Karachay-Balkar *See:* Turkic Languages.

Karaim *See:* Turkic Languages.

Karakalpak *See:* Turkic Languages.

Karuk *See:* Hokan Languages.

Kashaya *See:* Pomoan Languages.

Kashmiri

S Munshi, University of Texas at Austin, Austin, TX, USA

Kashmiri, known to its speakers as kə:śur/ko:śur or kə:śɨr zabān ('Kashmiri language'), is spoken by around 4 million people in India (*Ethnologue*, n.d.), primarily in the Kashmir valley (kəśi:r) and its surrounding regions in the state of Jammu and Kashmir (J & K). Sizeable population of Kashmiri speakers also live in other states of India. Approximately 0.1 million Kashmiri speakers have been reported to live in Pakistan (*Ethnologue*, n.d.). Kashmiri speakers are also found in various other countries. Major regional dialects of the language spoken in J & K state include: Standard Kashmiri (in and around Srinagar; used for educational and literary purposes), Kashtwāri/Kishtwāri (Kishtwār valley, southeast of Kashmir valley); Poguli (southern Banihal); Sirāji (several villages of district Doda); Rāmbani (Ramban and adjoining areas), and Bunjwali. Poguli and Sirāji are considerably influenced by Dogri (an Indo-Aryan language close to Punjabi (Panjabi)). Speakers of many of these dialects often argue for the separate existence of their native speech as full-fledged languages rather than dialects of Kashmiri. Information available on Kashmiri dialects is generally based on secondary sources.

There is a possibility of existence of more regional dialects in addition to the above-mentioned, but no detailed research on Kashmiri dialectology is available so far. Besides the major regional dialects, there are conspicuous differences between urban and rural Kashmiri on the one hand (in terms of accent, phonology, and lexicon; rural Kashmiri has preserved many archaic forms not used in urban speech) and between Muslim and Hindu Kashmiri on the other (mainly in terms of borrowings; while Hindus use more Sanskrit loans, Muslims commonly borrow from Persian (Farsi, Western) and Arabic for the corresponding words. The terms 'Persianized' and 'Sanskritized' Kashmiri are sometimes used to refer to these social dialects).

Historical Development

Kashmiri has been classified along with a number of languages grouped under the title 'Dardic.' Dardic languages are spoken in the extreme north of India and northwestern Pakistan, extending into Afghanistan. There has been considerable debate over the classification of Dardic languages as to whether they are a third branch of Indo-Iranian language family (other two being Indo-Aryan and Iranian), or (at least, some of them) are of pure Indo-Aryan origin. Dardic languages have preserved many archaic Indo-Iranian features otherwise lost in the modern Indo-Aryan languages. They have also developed certain features not found in other Indo-Iranian languages. Nevertheless, the term 'Dardic' constitutes a geographical convention rather than a linguistic expression. Like other Dardic languages, Kashmiri has similarities with both Indo-Aryan as well as Iranian. After continuing debates over a long period of time, many linguists have agreed upon an Indo-Aryan origin for Kashmiri. The term 'Dardic,' however, has gained much popularity, and is still used in view of the regional peculiarities shared by Kashmiri and other languages of the group.

Kashmiri belongs to the North-Western group of the Middle Indo-Aryan (MIA) languages/dialects, which includes several Dardic languages (e.g., Shina, Khowar, Torwali), Punjabi, Sindhi, and Lahnda (Panjabi, Western). One of the characteristic features with respect to the phonological system of the North-Western group is the retention of certain features lost elsewhere. In many modern Indo-Aryan languages, Old Indo-Aryan (OIA) sibilants – *ś* (palatal), *s* (dental) and *ṣ* (retroflex) – merged into dental *s*. Kashmiri retains two sibilants, the palatal *ś* and the dental *s*. Like other Dardic languages, Kashmiri also retains the consonantal component *r* in the derivatives of the OIA syllabic *r̥*, which had a number of reflexes in MIA, viz., *a*, *i*, or *u*.

Kashmiri vocabulary can be broadly categorized into Kashmiri/Dardic, Sanskrit, Punjabi, Hindi/Urdu, Persian, and Arabic origins. Kashmiri occupies a special position in the Dardic group, being probably the only Dardic language that has a written literature dating back to the early 13th century, a writing script of its own, and the largest number of speakers among the Dardic languages. An important part of Kashmiri is its rich tradition of oral and written literature. Some of the very famous genres of Kashmiri oral literature and folklore are *rov*, *vanɨvun*, *c^hakɨr*, *ladi:-śah*, and *lukɨ-pə:t^hɨr*.

Writing Systems

Originally, Kashmiri was written in the Shāradā script, an ancient indigenous character of Kashmir. Shāradā is argued to be the predecessor of Devanāgri/Nāgri, which is built on the same system and corresponds with Shāradā letter-for-letter, although the letters

have considerably changed in form. Shāradā is closely associated with Tākri alphabet used for writing Punjabi, but, with a complete range of symbols for different vowels characteristic of Kashmiri. Its use is highly restricted to a handful of Hindu priests in writing *za:tuk/Janam-patri* ('horoscope'). The most popularly employed and officially-recognized script in current use is a modification of Perso-Arabic (Nastālīq) script. Devanāgari (again with modifications to cater to the specific requirements of the Kashmiri phonemic inventory) is also used and is popular among Hindus. Tākri (Kashtwāri and some dialects of the adjoining areas) and Roman scripts have also been employed but these have failed to gain recognition.

Phonological Characteristics

Both open and closed syllables are permitted in Kashmiri. Closed syllables, however, are preferred to open syllables. In rapid speech, in polysyllabic words with a sequence of adjacent CV syllables, speakers may drop medial vowels in favor of closed syllable structure. Final vowels are often deleted. Clusters comprising two consonants are quite common but only specific sequences can form a cluster in a particular position. Initial clusters are restricted to the type *Cr-* where the first consonant of the cluster is an obstruent. Final clusters are comprised of a homorganic nasal followed by an obstruent. The distribution of stress in Standard Kashmiri is influenced by a complex interplay of quantitative, positional, and rhythmic constraints. Primary stress appears on the word-initial syllable, which is always stressed. Stress occurs on every syllable containing a long vowel – CV:(C).

Kashmiri maintains the basic OIA pattern of five articulatory positions along with features of voice (e.g., *k* vs. *g*) and aspiration (e.g., *k* vs. k^h). Its consonant system has survived with various alterations, losses, and additions. Some of the characteristic changes are those involving the loss of aspiration of voiced plosives, change of MIA palatal affricates to those of the corresponding dental affricates, word-final aspiration of voiceless stops, and fricative weakening/lenition ($s > h$) in certain environments. There are 27 consonantal phonemes in Kashmiri that have evolved from the Old/Middle Indo-Aryan phonological system. In addition, Kashmiri has also adopted a number of consonants from the Persian and Arabic phonemic inventory. The latter include the labio-dental voiceless fricative (*f*), voiceless and voiced velar fricatives (*x* and *γ*), and uvular and glottalic stops (*q* and *ʔ*). They are found only in Persian/Arabic borrowings and are used only by literate Kashmiris in Standard Kashmiri, especially in formal speech. In informal speech and among the illiterate population, these are replaced by p^h, k^h, *g*, *k*, and *a* respectively.

Like most of the New Indo-Aryan (NIA) languages, Kashmiri vowels are subject to various phonological operations that are not only regular but also extended over a larger domain than most of the other IA languages. A significant number of changes take place in accordance with the nature and position of the vowels in a particular linguistic domain (syllable/morpheme/word). Various such changes include: vowel harmony, *svarabhakti* (vowel epenthesis), consonantal assimilation, and final vowel deletion. Kashmiri has a 16-vowel system consisting of front vowels /*i*, *i:*, *e*, *e:*, *ɛ*/, central vowels /*ɨ*, *ɨ:*, *ə*, *ə:*, *a*, *a:*/, and back vowels /*u*, *u:*, *o*, *o:*, *ɔ*/ with three contrasts in height (high, mid, and low). Kashmiri phonemic inventory is distinct in the IA languages in having central vowels /*ɨ*, *ɨ:*, *ə*, *ə:*/ (absent in most NIA languages) and dental affricates.

Morphosyntax

Kashmiri, like other IA languages, is a postpositional language. However, its word order is unique among the IA languages. Unlike other IA languages, which are typically verb-final, Kashmiri is a V2 language. That is, the inflected verb occurs at the clause-second position. In sentences with a main verb and an inflected auxiliary verb, the main verb occurs sentence-finally. Thus, the basic word order is essentially SVO(V). This is true of the matrix as well as embedded clauses, and also of *yes-no* questions and questions where the *wh*-phrase is the syntactic 'subject.' In other question formations, the *wh*-phrase occupies the clause-second position, with the inflected verb occupying third position. In most environments, V2 is obligatory. However, in certain syntactic environments, such as correlatives and conditional clauses, V2 is optional. The sentence initial position is occupied by syntactic 'subject' or any other constituent ('topic'). Except for the fixed position of the verb (and *wh*-phrase in case of questions), word order is fairly flexible. V2, characteristic of Germanic languages, is a well-developed syntactic phenomenon in Kashmiri, a non-Germanic language.

Complementation in Kashmiri is observed by optionally inserting complementizer *ki/zi* in front of the embedded clause. Relative clauses are formed by a relative-correlative construction with a pre-nominal relative pronoun in the modifying/relative clause and a correlative pronoun in the main clause following the relative clause. Relative-correlative construction is a typical areal syntactic feature.

Kashmiri is a split-ergative language. Case morphology is more or less typically Indo-Aryan. Based on thematic roles and verb class, the only argument of an intransitive clause receives either a zero/nominative case or a dative case. The most agent-like argument of a transitive clause in perfective aspect receives an ergative case and the other argument (if any) receives a nominative or dative case depending on its thematic role and the verb class. Subjects of a few intransitive verbs may appear in ergative case (e.g., verbs like *as-un* 'to laugh', *vad-un* 'to weep/cry'). Case ending behaves like a postposition so that the noun phrase appears in oblique form (a typical Indo-Aryan syntactic feature).

Kashmiri verb phrase is rich in agreement. Both subject and object agreement markers may appear on the verb. Subject agreement, however, is blocked in dative/ergative constructions where the syntactic subject is in dative/ergative case.

One of the characteristic features of Kashmiri among the Indo-Aryan languages is its three-way (instead of the typical two-way) distinction of the demonstrative pronoun, viz., (1) proximate *yi/yim* 'this/these', (2) visible *hu/hum* 'that/they (masculine)' and *ho/homa* 'that/they (feminine)', and (3) invisible/remote *su/tim* 'that/they (masculine)' and *so/tima* 'that/they (feminine)'.

Sociolinguistics of Kashmiri

Not many native speakers of Kashmiri can read the language, irrespective of their educational background. For a long time, Kashmiri has not been taught in schools. According to the three-language policy of India, languages generally taught in the schools of J & K state are Urdu (a non-native language and the official language of the state), Hindi (official language of India), and English (second official language). There is a significant amount of prestige associated with Urdu, Hindi, and English in wider linguistic domains. These factors and increasing urbanization and globalization have played a significant role toward a gradual language loss with many Kashmiri speakers.

Sample (Srinagar/Standard Kashmiri)

(1)	ga:ś$_i$-an	von	sali:m$_j$-as	(zi/ki)
	Gasha-ERG.	say.PAST.PERF.	Salim-DAT.	COMP
	agar	a:si-hund$_k$	mo:l$_k$	
	if	Asi-Gen.M.SG.	father	
	gar-i	o:s$_k$	tam-is	
	home-LOC.	be.PAST.M.SG.	3SG.OBL.-DAT.	
	van-un	ma:j-i	samk^h-un	
	tell-Infin.	mother.OBL.-DAT.	meet-INFIN.	

'Gasha$_i$ told Salim$_j$ (that) if Asi's father$_k$ was home, he$_j$ should tell him$_k$ to meet (X's$_{i/j/k/l}$) mother'

See also: Areal Linguistics; Dardic; Harmony; India: Language Situation; Indo–Aryan Languages; Indo–Iranian; Lahnda; Pakistan: Language Situation; Punjabi; Sindhi.

Bibliography

Bashir E (2003). 'Dardic.' In Cardona G & Jain D (eds.) *The Indo-Aryan languages*. London/New York: Routledge, Taylor & Francis Group. 818–894.

Ethnologue (n.d.) 'Kashmiri'. *Languages of the World*, 15th edition.

Grierson G A (1906). *The Piśaca languages of North-Western India*. London: The Royal Asiatic Society.

Grierson G A (1919). *The linguistic survey of India VIII, Part II*. Calcutta: Superintendent Government Printing, India.

Koul O N (2003). 'Kashmiri.' In Cardona G & Jain D (eds.) *The Indo-Aryan languages*. London/New York: Routledge, Taylor & Francis Group. 895–952.

Koul O N & Hook P E (eds.) (1984). *Aspects of Kashmiri linguistics*. New Delhi: Bahri Publications Pvt. Ltd.

Masica C P (1991). *The Indo-Aryan languages*. Cambridge: Cambridge University Press.

Morgenstierne G (1941). 'The Phonology of Kashmiri.' In *Acta Orientalia 19(1)*. Reprinted in 1973 in *Irano-Dardica*. Wiesbaden: Dr. Ludwig Reichert Verlag. 79–99.

Kachru B B (1969). 'Kashmiri and other Dardic Languages.' In Sebeok T A (ed.) *Current trends in linguistics 5: Linguistics in South Asia*. Paris/The Hague: Mouton. 284–306.

Relevant Website

http://www.ethnologue.com/show_language.asp?code=kas

Kate Hermansz, Lambert ten (1674–1731)

J Noordegraaf, Vrije Universiteit Amsterdam

Lambert ten Kate, who is considered to be one of the founding fathers of Germanic linguistics, was born of Mennonite parents in Amsterdam. He became a partner in his father's corn trade but resigned in 1706. "Being wealthy enough," as his first biographer has it, the versatile ten Kate "laid aside Commerce, in order to cultivate his Studies," which included various fields of both arts and sciences.

In 1707–1710 ten Kate composed his short *Gemeenschap tussen de Gottische spraeke en de Nederduytsche* ('The relationship between the Gothic and Dutch languages'). The insight that the conjugation of verbs in Dutch and Gothic follows the same pattern prompted ten Kate to divide the Gothic verbs into six classes. It is in the verbal conjugation system that ten Kate recognized the regularity of vowel alternation, which eventually led him to the discovery of what Grimm would later call 'Ablaut.' In all respects the *Gemeenschap* has proved to be more valuable and more fruitful than the work of his English predecessor, Hickes.

The years 1710–1723 saw the full development of these and other interesting linguistic observations into the main body of his *Aenleiding tot de kennisse van het verhevene deel der Nederduitsche sprake* ('Introduction to the sublime part of the Dutch language,' i.e., its etymology). Among other things, the classification of the Gothic verbs was elaborated here and applied to all the Germanic languages he was acquainted with. Ten Kate provided proof that the 'irregular' verbs were subject to rules that obtained not only in Gothic, but were valid for all branches of Germanic. The second volume of the *Aenleiding* includes two long alphabetically arranged listings of Germanic words derived from strong verb roots. As ten Kate deemed the verb to be the very first part of speech, and consequently the uppermost in the hierarchy of the *partes orationis,* this specimen of a *lexicon etymologicum* could thus be considered as a lexicon of root words.

Ten Kate's main concern was to find regularity in language, taking great care to test the rules he had discovered against an extensive corpus of linguistic data. Abandoning the techniques of ancient etymology, the Amsterdam private scholar described and justified his own research methods very carefully. Thus, he actively contributed to the empirical turn in the intellectual climate of the Dutch Republic at the beginning of the eighteenth century. As such, his works are a fine illustration of the paradigm shift that took place in Dutch linguistics around 1700 under the influence of Newtonian thought. They stand out in the 18th century as a highly successful contribution to the study of historical linguistics, in terms of both its contents and its methodology.

See also: Grimm, Jacob Ludwig Carl (1785–1863).

Bibliography

Kate Hermansz L ten (1710). *Gemeenschap tussen de Gottische spraeke en de Nederduytsche.* Amsterdam: Jan Rieuwertsz. [Reprinted 2002, Amsterdam: Stichting Neerlandistiek VU.]

Kate Hermansz L ten (1723). *Aenleiding tot de kennisse van het verhevene deel der Nederduitsche sprake* (2 vols). Amsterdam: Rudolph en Gerard Wetstein. [Reprinted 2001, Alphen aan den Rijn: Canaletto.]

Kate Hermansz L ten (1732). *The beau ideal, by the late ingenious and learned Hollander Lambert Hermanson Ten Kate, translated from the original French by James Christopher Le Blon.* London: Bettenham.

Le Blon J C (1732). 'The translator's preface.' In ten Kate 1732/2001, [i]–iv.

Noordegraaf J (2003). '"Let us now praise famous men": underlying conceptions in the works of Lambert ten Kate.' In Dutz K (ed.) *Später Mittag. Vermischte Anmerkungen zur Metahistoriographie.* Münster: Nodus. 175–190.

Noordegraaf J & van der Wal M J (2001). 'Lambert ten Kate and Linguistics.' In ten Kate 1723/2001 I, 2–32 [with extensive bibliography].

Rutten G J (2004). 'Lambert ten Kate and Justus-Georg Schottelius: theoretical similarities between Dutch and German early modern linguistics.' *Historiographia Linguistica 31*, 277–296.

Salverda R (2001). 'Newtonian linguistics: the contribution of Lambert ten Kate (1674–1731) to the study of language.' In Davies M C, Flood J & Yeandle D N (eds.) '*Proper words in proper places*': *studies in lexicology and lexicography in honour of William Jervis Jones.* Stuttgart: Hans-Dieter Heinz Akademischer Verlag. 115–132.

Katre, Sumitra Mangesh (1906–1998)

F Britto, Sophia University, Tokyo, Japan

An outstanding scholar, educator, and administrator, Sumitra Mangesh Katre was born in Honnavar, Karnataka, on April 11, 1906. After his schooling in Mangalore, he obtained a B.A. (Hons) in mathematics and Sanskrit from Madras University (1928). Subsequently, he went to the School of Oriental and African Studies, London, and obtained a Ph.D. in linguistics (1931). On his return to India, he taught at several colleges before entering the Deccan College Postgraduate and Research Institute, Pune, as Professor of Indo-European Philology in 1939. Three years later, he became the director of the institute and served in that capacity until his retirement in 1971. He then moved to the United States as a resident and taught in the Department of Oriental and African Languages and Literature at the University of Texas, Austin (1971–1976). He died on October 21, 1998, in San Jose, California.

Katre wrote 12 books and approximately 250 articles and edited several tomes on linguistics. His most noteworthy publications relate to Sanskrit manuscripts (Katre, 1949–1955), Paninian studies (Katre, 1967, 1987), Paninian lexicography (Katre, 1968–1969, 1971, 1981), Prakrit and Indo-Aryan studies (Katre, 1944, 1964, 1968), and the formation of Konkani (Katre, 1966). His *Introduction to Indian textual criticism* (1954) and *Lexicography* (1965) are highly practical textbook-like volumes.

Katre is perhaps best remembered for his dedication to the spread of linguistics in India. He pioneered the teaching of linguistics in Indian universities by actively initiating programs, influencing authorities, and educating future linguists. The Language Project, of which he was the director from 1954 to 1960, held linguistic conferences for the first time in India and conducted summer and fall schools of linguistics for the formation of linguists. Due to the Project, illustrious linguists from overseas visited India and promising Indians went overseas for linguistic training. Katre also contributed to the spread of linguistics by his active involvement in several linguistic journals, such as *New Indian Antiquary*, *The Oriental Digest*, *Indian Linguistics*, *The Bulletin of the Deccan College*, and *The Indo-Iranian Journal*, serving as an editor, board member, or founder. He gave new life to the Linguistic Society of India by having it moved from Calcutta to Deccan College, Pune, and making it the preeminent all-India organization of linguists.

A monumental project that Khatre conceived and headed for 20 years (1951–1971) is the *Encyclopaedic dictionary of Sanskrit on historical principles*. This dictionary was more than 25 years in preparation, involving excerpts from 1500 books from the Vedic period to the 18th century, and based on a data bank of 11 million words. Although the original project has been modified, several volumes have been published, and when completed, with approximately 20 000 pages, it is expected to be the largest dictionary ever made.

Katre held many administrative and consultative positions in leading academic institutions and organizations, such as the Bhandarkar Oriental Research Institute, the University Grants Commission, and the Central Institute for Indian Languages. He represented the government of India at many international conferences, such as the Congress of Linguistics in London (1952), the Congress of Linguistics in Oslo (1957), and the Munich Session of the International Congress of Orientalists (1957).

He delivered the *Wilson philological lectures* at Bombay University (1941) and the *Pratibha Devi memorial lectures* at Gauhati University (1957). He was honored with two felicitation volumes: *Indian Linguistics* volumes 29 and 30 (1968–1969) and the *Bulletin of the Deccan College Postgraduate and Research Institute* volumes 51 and 52 (1991–1992).

See also: Sanskrit.

Bibliography

An encyclopaedic dictionary of Sanskrit on historical principles (1976–2002). Pune, India: Deccan College Postgraduate and Research Institute.

Anonymous (1969). 'Bibliography of Dr. S. M. Katre's writings.' *Indian Linguistics 30* (Katre Felicitation volume, part II), xiii–xvi.

Anonymous (1997). "Katre, S. M." In *Dravidian encyclopaedia: Vol. 3. Language and literature*. Thiruvananthapuram, India: International School of Dravidian Linguistics. 340.

Katre S M (1944). *Some problems of historical linguistics in Indo-Aryan (Wilson philological lectures 1940–41)*. Bombay: University of Bombay.

Katre S M (1949–1955). *Descriptive catalogue of Govt. Mss. at the Bhandarkar Oriental Research Institute: Vol. 9. Vedanta* (in 3 parts). Pune, India: Bhandarkar Oriental Research Institute.

Katre S M (1954). *Introduction to Indian textual criticism*. Pune, India: Deccan College Postgraduate and Research Institute.

Katre S M (1964). *Prakrit languages and their contribution to Indian culture.* Pune, India: Deccan College Postgraduate and Research Institute.
Katre S M (1965). *Lexicography: a course of three special lectures delivered at the Linguistics Department of Annamalai University.* Annamalai University Publications in Linguistics no. 6. Annamalainagar, India: Annamalai University.
Katre S M (1966). *The formation of Konkani.* Pune, India: Deccan College Postgraduate and Research Institute.
Katre S M (1967). *Paninian studies I.* Pune, India: Deccan College Postgraduate and Research Institute.
Katre S M (1968). *Problems of reconstruction in Indo-Aryan.* Simla, India: Institute of Advanced Study.
Katre S M (1968–1969). *Dictionary of Panini* (3 vols). Pune, India: Deccan College Postgraduate and Research Institute.
Katre S M (1971). *Dictionary of Panini: Ganapatha.* Pune, India: Deccan College Postgraduate and Research Institute.
Katre S M (1981). *A glossary of grammatical elements and operations in Astadhyayi.* Mysore, India: Central Institute of Indian Languages.
Katre S M (1987). *Astadhyayi of Panini in Roman transliteration.* Austin: University of Texas.
Patyal H C (1999). 'Obituary: Sumitra Mangesh Katre.' *Indo-Iranian Journal 42,* v–vi.

Kātyāyana (3rd Century B.C.)

S Shukla, Georgetown University, Washington, DC, USA

Pāṇini, Kātyāyana, and Patañjali belong to the beginning of the grammatical tradition in India. We have no records of linguistics prior to Pāṇini, whose grammar of the Sanskrit language, *Aṣṭādhyāyī*, has commanded the respect of linguists everywhere. But the quest for economy inspired by the oral tradition eventually made the task of comprehending it enormously complicated and it became intelligible only when studied with a commentary.

The major commentaries on Pāṇini are the *Vārttikas* (Explanatory rules) of Kātyāyana, which examine approximately one-third of Pāṇini's *sūtras*, and the *Mahābhāṣya* (The great commentary) of Patañjali, which is mainly concerned with explaining the *vārttikas* of Kātyāyana.

Patañjali's date (2nd century B.C.) is known through the references to the Shunga king Pushyamitra and to the invasion of the Bactrian Greeks (Belvalkar, 1976: 27). A considerable interval must have elapsed between Kātyāyana and Patañjali and between Pāṇini and Kātyāyana. Accordingly, Pāṇini is believed to have lived in the 5th century B.C., which is in accordance with the native tradition that connects him with the and king of Magadha, and Kātyāyana is believed to have lived not long after Pāṇini, in the 3rd century B.C. (Burrow, 1973: 48; Belvalkar, 1976: 24).

Kātyāyana, a descendant of the sage Kati, is known to have written two other works, the *Vājasaneyi Prātiśākhya* and *Kātyāyana śrauta Sūtra* (Belvalkar, 1976: 24). In his *Mahābhāṣya*, Patañjali calls Kātyāyana a *dākṣiṇātya* 'Southerner' (Ray, 1963: 89; Belvalkar, 1976: 25). He has been also identified with Vara-ruci, the author of *Prākrita-Prakāśa*. The *vārttikas*, for which Kātyāyana is justly famous, have not survived separately in manuscript form. They occur in Patañjali's *Mahābhāṣya*, which is mainly concerned with explaining and criticizing the *vārttikas*.

When commenting on a Pāṇini's *sūtra* (grammatical rules), Kātyāyana either repeats it or refers to it by incorporating part of it in his *vārttika*. When cross-referencing his own *vārttikas*, Kātyāyana uses the phrase "*uktaṃ vā*" or "*uktaṃ pūrveṇa*" '(with regard to this) a statement has (already) been made by an earlier rule'. In his *vārttikas,* he often uses linguistic terms not used by Pāṇini, for example, *svara* 'vowel' (for Pāṇini's *ac*) and *vyañjana* 'consonant' (for Pāṇini's *hal*). It becomes clear from his *vārttikas* on Pāṇini 8.2.83 and 8.2.84 that Kātyāyana was commenting on a language spoken by him as his mother tongue (Vasu, 1962: 2: 1582). As an example of an operation of Kātyāyana's *vārttikas,* consider the following (Joshi and Roodbegen, 1981: 25; Vasu, 1982: 300).

- Pāṇni's *sūtra* (P2.3.61).

preṣyabruvorhaviṣo devatāsampradāne

"(In the Vedas the genitive case ending is used to express the sense of accusative) after (a word denoting) offering in connection with (the verbal bases) *preṣya-* 'to send forth' and *bru-* 'to declare,' when an oblation is made to a deity."

- Kātyāyana's rectification *vārttika* on P2.3.61.

haviṣo'prasthitasya

"(The Pāṇinian phrase *haviṣaḥ* should be changed to) *haviṣo'prasthitasya* 'an oblation which has not been held ready'."

- Patañjali's explanation in his *Mahābhāṣya*, accepting Kātyāyana's *vārttika*.

haviṣo'prasthitasyeti vaktavyaṃ

Katre S M (1964). *Prakrit languages and their contribution to Indian culture.* Pune, India: Deccan College Postgraduate and Research Institute.
Katre S M (1965). *Lexicography: a course of three special lectures delivered at the Linguistics Department of Annamalai University.* Annamalai University Publications in Linguistics no. 6. Annamalainagar, India: Annamalai University.
Katre S M (1966). *The formation of Konkani.* Pune, India: Deccan College Postgraduate and Research Institute.
Katre S M (1967). *Paninian studies I.* Pune, India: Deccan College Postgraduate and Research Institute.
Katre S M (1968). *Problems of reconstruction in Indo-Aryan.* Simla, India: Institute of Advanced Study.
Katre S M (1968–1969). *Dictionary of Panini* (3 vols). Pune, India: Deccan College Postgraduate and Research Institute.
Katre S M (1971). *Dictionary of Panini: Ganapatha.* Pune, India: Deccan College Postgraduate and Research Institute.
Katre S M (1981). *A glossary of grammatical elements and operations in Astadhyayi.* Mysore, India: Central Institute of Indian Languages.
Katre S M (1987). *Astadhyayi of Panini in Roman transliteration.* Austin: University of Texas.
Patyal H C (1999). 'Obituary: Sumitra Mangesh Katre.' *Indo-Iranian Journal 42,* v–vi.

Kātyāyana (3rd Century B.C.)

S Shukla, Georgetown University, Washington, DC, USA

Pāṇini, Kātyāyana, and Patañjali belong to the beginning of the grammatical tradition in India. We have no records of linguistics prior to Pāṇini, whose grammar of the Sanskrit language, *Aṣṭādhyāyī*, has commanded the respect of linguists everywhere. But the quest for economy inspired by the oral tradition eventually made the task of comprehending it enormously complicated and it became intelligible only when studied with a commentary.

The major commentaries on Pāṇini are the *Vārttikas* (Explanatory rules) of Kātyāyana, which examine approximately one-third of Pāṇini's *sūtras*, and the *Mahābhāṣya* (The great commentary) of Patañjali, which is mainly concerned with explaining the *vārttikas* of Kātyāyana.

Patañjali's date (2nd century B.C.) is known through the references to the Shunga king Pushyamitra and to the invasion of the Bactrian Greeks (Belvalkar, 1976: 27). A considerable interval must have elapsed between Kātyāyana and Patañjali and between Pāṇini and Kātyāyana. Accordingly, Pāṇini is believed to have lived in the 5th century B.C., which is in accordance with the native tradition that connects him with the and king of Magadha, and Kātyāyana is believed to have lived not long after Pāṇini, in the 3rd century B.C. (Burrow, 1973: 48; Belvalkar, 1976: 24).

Kātyāyana, a descendant of the sage Kati, is known to have written two other works, the *Vājasaneyi Prātiśākhya* and *Kātyāyana śrauta Sūtra* (Belvalkar, 1976: 24). In his *Mahābhāṣya*, Patañjali calls Kātyāyana a *dākṣiṇātya* 'Southerner' (Ray, 1963: 89; Belvalkar, 1976: 25). He has been also identified with Vara-ruci, the author of *Prākrita-Prakāśa*. The *vārttikas*, for which Kātyāyana is justly famous, have not survived separately in manuscript form. They occur in Patañjali's *Mahābhāṣya*, which is mainly concerned with explaining and criticizing the *vārttikas*.

When commenting on a Pāṇini's *sūtra* (grammatical rules), Kātyāyana either repeats it or refers to it by incorporating part of it in his *vārttika*. When cross-referencing his own *vārttikas*, Kātyāyana uses the phrase "*uktaṃ vā*" or "*uktaṃ pūrveṇa*" '(with regard to this) a statement has (already) been made by an earlier rule'. In his *vārttikas,* he often uses linguistic terms not used by Pāṇini, for example, *svara* 'vowel' (for Pāṇini's *ac*) and *vyañjana* 'consonant' (for Pāṇini's *hal*). It becomes clear from his *vārttikas* on Pāṇini 8.2.83 and 8.2.84 that Kātyāyana was commenting on a language spoken by him as his mother tongue (Vasu, 1962: 2: 1582). As an example of an operation of Kātyāyana's *vārttikas,* consider the following (Joshi and Roodbegen, 1981: 25; Vasu, 1982: 300).

- Pāṇni's *sūtra* (P2.3.61).

preṣyabruvorhaviṣo devatāsampradāne

"(In the Vedas the genitive case ending is used to express the sense of accusative) after (a word denoting) offering in connection with (the verbal bases) *preṣya-* 'to send forth' and *bru-* 'to declare,' when an oblation is made to a deity."

- Kātyāyana's rectification *vārttika* on P2.3.61.

haviṣo'prasthitasya

"(The Pāṇinian phrase *haviṣaḥ* should be changed to) *haviṣo'prasthitasya* 'an oblation which has not been held ready'."

- Patañjali's explanation in his *Mahābhāṣya*, accepting Kātyāyana's *vārttika*.

haviṣo'prasthitasyeti vaktavyaṃ

"The (statement) *haviṣo'prasthitasya* should be made. (For instance) *indrāgnibhyām chāgam havirvapāṃ medaḥ prasthitaṃ preṣya* 'Send forth to *Indra* and *Agni* the oblation (consisting of the organs of) a goat, marrow, (and its) fat which are held in readiness'."

In this example given by Patañjali, the word *haviṣ* is qualified by *prasthita* 'held in readiness', and the genitive ending is therefore not used. Instead, the words *haviṣ*, *chāga*, *vapa*, and *medas* occur with accusative endings.

There are some *vārttikas* of Kātyāyana that are, however, overruled by Patañjali (Vasu, 1962: 1: 300–301; Joshi and Roodbegen, 1981: 25), as in the following example.

- Pāṇini's *sūtra* (P2.3.62).

caturthyarthe bahulaṃ chandasi

"The sixth (genitive case ending is) often (used) to express the sense of the fourth (dative case ending) in the *Veda*." (For example, *puruṣmṛgas candramasaḥ* 'the male antelope to the very moon' in which the genitive *candramsaḥ* (not the dative *candramase*) is used.)

- Kātyāyana's *vārttika* offers an addition to P2.3.62.

ṣaṣṭhyarthe caturthīvacanaṃ

"The fourth (dative) case ending should be prescribed to express the sense of the sixth (genitive case ending in the Veda)."

- Patañjali, citing Kātyāyana's *vārttika,* illustrates it with this example.

yā kharveṇa pibati tasyai kharvas. tasyā it prāpte...

"'to her who drinks from a damaged vessel a defective (son is born)'. (Here) when (the genitive form) *tasyāh* (genitive 'to her') is about to apply (the dative form *tasyai* 'for her' is used)." But then Patañjali comments that the *vārttika* is not needed if the *sūtra* P2.3.62 is split into two:

caturthyarthī

"the fourth case ending (is used to express the sense of the sixth case ending)" and

tato'rthe bahulaṃ chandasīti

"to express the sense (of the fourth case ending the sixth case ending is) often used in the Veda."

Hence, Patañjali sees that once Pāṇini's *sūtra* P2.3.62 is understood as having two constituents, both the genitive and dative uses are explained by it and Kātyāyana's *vārttika*, suggesting an addition, is not needed.

Commenting on Kātyāyana's *vārttikas*, the celebrated 18th-century grammarian Nāgojibhaṭṭa said: "*vārttikamiti. sūtre'nuktaduruktacintākaratvaṃ vārttikatvam*" ("The characteristic feature of a *vārttika* is criticism in regard to that which is omitted or imperfectly expressed in a *sūtra* of Pāṇini") (Kielhorn, 1985: 2). It must be mentioned that a critical look at the *vārttikas* shows Kātyāyana to be a judicious admirer of Pāṇni, one who dispassionately examines Pāṇini's *sūtras*; considers objections that have been actually raised or that might be raised; and corrects them, offers additions to them, or advises that they be abandoned only when no other course is left open (Kielhorn, 1963: 47).

See also: Nagojibhatta (d. 1755); Panini; Patanjali (2nd Century B.C.).

Bibliography

Abhyankar K V (ed.) (1960). *Paribhāṣenduśekhara of Nāgeśa.* Kielhorn P (trans. and notes). Poona, India: Bhandarkar Oriental Research Institute.

Abhyankar K V & Limaye V P (eds.) (1965). *Vākya-Padīya of Bhartṛhari.* Poona, India: University of Poona.

Agnihotri P (1963). *Patañjalikālīna Bhārata.* Patna, India: Bihar-Rashtrabhasha-Parishada.

Agrawal V S (1963). *India as Known to Pāṇini.* Varanasi, India: Prithivi Prakashan.

Allen W S (1953). *Phonetics in ancient India.* London: Oxford University Press.

Belvalkar S K (1976). *An Account of different existing systems of Sanskrit Grammar* (reprint). Delhi, India: Bhartiya Vidya Prakashan.

Bhartṛhari (1922). *Subhaṣita Triśati.* Bombay, India: Nirnaya Sagar Press.

Bloomfield L (1927). 'On some rules of Panini.' *Journal of the American Oriental Society 47,* 61–70.

Bloomfield L (1929). 'Review of Leibich.' *Language 5,* 267–275.

Bloomfield L (1931). *Language.* New York: Henry Holt.

Brough J (1951). 'Theories of general linguistics in the Sanskrit grammarians.' *Transaction of the Philological Society 27(46),* 402–414.

Burrow T (1973). *The Sanskrit language.* London: Faber and Faber.

Cardona G (1997). *Pāṇini: his work and its tradition.* Delhi, India: Motilal Banarasidass.

Chaitanya K (1962). *A new history of Sanskrit literature.* Westport, CT: Greenwood Press.

Joshi B (ed.) (1942). *The Siddhānta Kaumudī with the Tattvabodhinī commentary.* Bombay, India: Nirnaya Sagar Press.

Joshi S D & Roodbergen J A F (1981). *Patañjali's Vyākaraṇa-Mahābhāṣya.* Pune, India: University of Poona.

Keith A B (1948). *A history of Sanskrit literature.* London: Oxford University Press.

Kielhorn F (1963). *Kātyāyana and Patañjali* (reprint). Varanasi, India: Indological Book House.
Kielhorn F (1985). *Paribhāṣenduśekhara of Nāgojibhaṭṭa* (reprint). Delhi, India: Parimal Publications.
Kudalal S & Shastri R (eds.) (1912). *Patañjali's Vyākaraṇa Mhābhāṣya with Kaiyaṭa's Pradīpa and Nāgeś's Udyota.* Bombay, India: Nirnaya-Sagar Press.
Macdonnell A A (1927). *India's past.* Oxford: The Clarendon Press.
Mishra V N (1966). *The descriptive technique of Pāṇini: An introduction.* The Hague: Mouton.
Ray K R (1963). *Patañjali's Mahābhāṣya-Vyākaraṇa.* Calcutta, India: Sanskrit Pustak Bhandar.
Robins R H (1979). *A short history of linguistics* (2nd edn.). London: Longman.
Sharma R N (1987). *The Aṣṭādhyāyī of Pāṇini* (vol. 1). Delhi, India: Munshiram Manoharlal Publishers.
Singh P (1987). *Paribhāṣenduśekhara.* Varanasi, India: Indu Prakashan.
Stall F (1972). *A reader on the Sanskrit grammarians.* Cambridge, MA: MIT Press.
Varma S (1961). *Critical studies in the phonetic observations of Indian grammarians.* Delhi, Inida: Munshi Ram Manohar Lal.
Vasu S K (1962). *The Aṣṭādhyāyī of Pāṇini* (2 vols) (reprint). Delhi, India: Motilal Banarasidass.
Vasu S K (1982). *The Siddhānta Kaumudī of Bhaṭṭoji Dīkṣita* (2 vols) (reprint). Delhi, India: Motilal Banarasidass.

Katz, Jerrold J. (1932–2002)

D T Langendoen, University of Arizona, Tucson, AZ, USA

Jerrold Katz's Career

Katz received a Ph.D. in Philosophy from Princeton University in 1960. In 1961 he became Research Associate in Linguistics at the Massachusetts Institute of Technology, and he was appointed Assistant Professor of Philosophy there in 1963. He quickly rose through the ranks at MIT, becoming Professor of Philosophy in 1969. He was appointed Distinguished Professor of Philosophy and Linguistics at the Graduate Center of the City University of New York in 1975, a position he held until his death. He wrote or coauthored 11 books and over 75 articles on philosophy and linguistics, and edited or co-edited three books.

Semantics in Generative Grammar

Katz started out with the goal "to try to understand meaning in natural language using some form of Fregean intensionalism to provide a semantic theory for Chomsky's theory of generative grammar, which ... at the time had no semantic theory" (Katz, 2004: ix). He published a seminal paper with Jerry Fodor (Katz and Fodor, 1963) that proposed that semantics is a component of generative grammar, and that semantic interpretations (senses) are assigned to syntactic structures by projection rules that have some of the formal properties of transformations. In addition, they proposed that the meanings of lexical items and of the structures containing them are expressed using combinations of semantic markers that are "intended to reflect, in their formal structure, the structure of the concepts they represent" (Katz, 1972: 101). A year later Katz published a book with Paul Postal (Katz and Postal, 1964) that proposed that deep structure (underlying syntactic structure) fully determines meaning (equivalently, that syntactic transformations are meaning-preserving). This idea, along with several others that were developed around the same time, was incorporated into the standard theory of generative grammar formulated in Noam Chomsky (1965), in which deep structures are mapped (1) by projection rules to semantic interpretations and (2) by syntactic transformations to surface structures, and ultimately to phonetic interpretations.

Katz (1972) provided the most complete account of his semantic theory; two later revisions were his elaboration of the theory of illocutionary force (Katz, 1977) and his revised treatment of presupposition (Katz and Langendoen, 1976; Katz, 1979). The book also provides a vigorous defense of his views against both generative semantics and the variety of interpretive semantics found in the extended standard theory.

Language as an Abstract Object

Katz was an early champion of Chomskyan mentalism (or psychologism or conceptualism) in linguistics (Katz, 1964), according to which linguistic theory is identified with the theory of knowledge of language. However, not long after collaborating with Chomsky in a defense of mentalism (Chomsky and Katz, 1975),

Kielhorn F (1963). *Kātyāyana and Patañjali* (reprint). Varanasi, India: Indological Book House.
Kielhorn F (1985). *Paribhāṣenduśekhara of Nāgojibhaṭṭa* (reprint). Delhi, India: Parimal Publications.
Kudalal S & Shastri R (eds.) (1912). *Patañjali's Vyākaraṇa Mhābhāṣya with Kaiyaṭa's Pradīpa and Nāgeś's Udyota.* Bombay, India: Nirnaya-Sagar Press.
Macdonnell A A (1927). *India's past.* Oxford: The Clarendon Press.
Mishra V N (1966). *The descriptive technique of Pāṇini: An introduction.* The Hague: Mouton.
Ray K R (1963). *Patañjali's Mahābhāṣya-Vyākaraṇa.* Calcutta, India: Sanskrit Pustak Bhandar.
Robins R H (1979). *A short history of linguistics* (2nd edn.). London: Longman.
Sharma R N (1987). *The Aṣṭādhyāyī of Pāṇini* (vol. 1). Delhi, India: Munshiram Manoharlal Publishers.
Singh P (1987). *Paribhāṣenduśekhara.* Varanasi, India: Indu Prakashan.
Stall F (1972). *A reader on the Sanskrit grammarians.* Cambridge, MA: MIT Press.
Varma S (1961). *Critical studies in the phonetic observations of Indian grammarians.* Delhi, Inida: Munshi Ram Manohar Lal.
Vasu S K (1962). *The Aṣṭādhyāyī of Pāṇini* (2 vols) (reprint). Delhi, India: Motilal Banarasidass.
Vasu S K (1982). *The Siddhānta Kaumudī of Bhaṭṭoji Dīkṣita* (2 vols) (reprint). Delhi, India: Motilal Banarasidass.

Katz, Jerrold J. (1932–2002)

D T Langendoen, University of Arizona, Tucson, AZ, USA

Jerrold Katz's Career

Katz received a Ph.D. in Philosophy from Princeton University in 1960. In 1961 he became Research Associate in Linguistics at the Massachusetts Institute of Technology, and he was appointed Assistant Professor of Philosophy there in 1963. He quickly rose through the ranks at MIT, becoming Professor of Philosophy in 1969. He was appointed Distinguished Professor of Philosophy and Linguistics at the Graduate Center of the City University of New York in 1975, a position he held until his death. He wrote or coauthored 11 books and over 75 articles on philosophy and linguistics, and edited or co-edited three books.

Semantics in Generative Grammar

Katz started out with the goal "to try to understand meaning in natural language using some form of Fregean intensionalism to provide a semantic theory for Chomsky's theory of generative grammar, which ... at the time had no semantic theory" (Katz, 2004: ix). He published a seminal paper with Jerry Fodor (Katz and Fodor, 1963) that proposed that semantics is a component of generative grammar, and that semantic interpretations (senses) are assigned to syntactic structures by projection rules that have some of the formal properties of transformations. In addition, they proposed that the meanings of lexical items and of the structures containing them are expressed using combinations of semantic markers that are "intended to reflect, in their formal structure, the structure of the concepts they represent" (Katz, 1972: 101). A year later Katz published a book with Paul Postal (Katz and Postal, 1964) that proposed that deep structure (underlying syntactic structure) fully determines meaning (equivalently, that syntactic transformations are meaning-preserving). This idea, along with several others that were developed around the same time, was incorporated into the standard theory of generative grammar formulated in Noam Chomsky (1965), in which deep structures are mapped (1) by projection rules to semantic interpretations and (2) by syntactic transformations to surface structures, and ultimately to phonetic interpretations.

Katz (1972) provided the most complete account of his semantic theory; two later revisions were his elaboration of the theory of illocutionary force (Katz, 1977) and his revised treatment of presupposition (Katz and Langendoen, 1976; Katz, 1979). The book also provides a vigorous defense of his views against both generative semantics and the variety of interpretive semantics found in the extended standard theory.

Language as an Abstract Object

Katz was an early champion of Chomskyan mentalism (or psychologism or conceptualism) in linguistics (Katz, 1964), according to which linguistic theory is identified with the theory of knowledge of language. However, not long after collaborating with Chomsky in a defense of mentalism (Chomsky and Katz, 1975),

Katz rejected mentalism in favor of (Platonic) realism, according to which languages are abstract objects, and the proper theory of language is about those objects themselves, not our knowledge of them (Katz, 1981). He described this change as follows (Katz, 2004: x):

> I had been wondering about how well Frege's realism about senses, to which I was committed, squared with Chomsky's psychologism about language, to which I was also committed. I reached the conclusion that . . . a theory of abstract senses could not be fitted into a theory of concrete syntactic structures in the human mind. My solution was to adopt a realist view of grammar as a whole, a move that seemed the right choice in light of the fact that the words and sentences that grammars are theories of are plausibly regarded as types and hence as abstract objects.

In moving from mentalism to realism, Katz saw himself as extending Chomsky's revolution that overthrew empiricism in favor of rationalism, and nominalism in favor of mentalism. A fully rationalist theory of language demands that it be realist, not mentalist (Katz, 1996; see also the papers in Katz, 1985). Katz's philosophy of mathematics was similarly realist and rationalist (Katz, 1998).

Sense, Reference, and Logic

Sense is the central construct of Katz's semantic theory, in which the meaning of an expression is the set of its senses. Katz defined sense as "that aspect of the grammatical structure of sentences that is responsible for their sense properties and relations" (Katz, 2004: 17). Sense is both decompositional (words and morphemes have internal structure in the form of semantic markers) and compositional (semantic markers are combined in accordance with their syntactic structure and projection principles). Katz did not provide a complete account of sense; for example, he did not list the primitive semantic markers. Rather, he gave many examples of senses of lexical items, such as color adjectives and English verbs like *chase* and *gossip*, and statements of particular projection rules for combining them. He used the framework to address such classical philosophical problems as the color-incompatibility problem (Katz, 2004: 152–165).

In later work, Katz asserted both that sense does not determine reference and that the theory of sense is not reducible to the theory of reference, unlike Frege's account, as interpreted by Rudolf Carnap, as functions from possible worlds to extensions in them (Katz, 2004: 21). Katz had earlier affirmed, following Frege, that sense determines reference (Katz, 1972: 240). But he went on to develop a maximally 'thin' notion of sense, according to which sense merely 'mediates' reference (Katz, 1988, 2004: 47). Katz's reformulation of the relation of sense to reference was intended to put intensionalist semantics on a firm foundation, enabling it to deal effectively with problems that Quine and others have raised, such as the indeterminacy of translation, the analytic-synthetic distinction, and the meaning of proper names and natural-kind terms. Katz also broke with the Frege-Carnap tradition by holding that senses are not represented by formulas in first-order predicate logic, which he maintained is insufficiently restrictive for linguistic semantics. His own theory of sense is mereological: semantic entailment is determined by the part-whole relations among senses, according to the 'beams in the house' analogy of Kant, rather than the 'plant in the seed' analogy of Frege.

Overall, Katz's work is distinguished by its bold independence of thought and its focus on foundational questions in philosophy of language, linguistics, and mathematics.

See also: Carnap, Rudolf (1891–1970); Chomsky, Noam (b. 1928); Fodor, Jerry (b. 1935); Frege, Gottlob (1848–1925); Kant, Immanuel (1724–1804); Mentalese; Philosophy of Mind; Quine, Willard van Orman (1908–2000).

Bibliography

Chomsky N (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky N & Katz J J (1975). 'On innateness: A reply to Cooper.' *The Philosophical Review 84,* 70–86.

Katz J J (1964). 'Mentalism in linguistics.' *Language 40,* 124–137.

Katz J J (1972). *Semantic theory.* New York: Harper & Row.

Katz J J (1977). *Propositional structure and illocutionary force: a study of the contribution of sentence meaning to speech acts.* New York: T. Y. Crowell.

Katz J J (1979). 'A solution to the projection problem for presupposition.' In Oh C-K (ed.) *Syntax and semantics, vol. 11: presupposition.* London: Academic Press. 91–126.

Katz J J (1981). *Language and other abstract objects.* Totowa, NJ: Rowman & Littlefield, and Oxford: Blackwell.

Katz J J (ed.) (1985). *The philosophy of linguistics.* Oxford: Oxford University Press.

Katz J J (1988). 'The refutation of indeterminacy.' *The Journal of Philosophy 85,* 227–252.

Katz J J (1996). 'The unfinished Chomskyan revolution.' *Mind and Language 11,* 270–294.

Katz J J (1998). *Realistic rationalism.* Cambridge, MA: MIT Press.

Katz J J (2004). *Sense, reference, and philosophy.* New York: Oxford University Press.
Katz J J & Fodor J A (1963). 'The structure of a semantic theory.' *Language 39,* 170–210.
Katz J J & Langendoen D T (1976). 'Pragmatics and presupposition.' *Language 52,* 1–17.
Katz J J & Postal P M (1964). *An integrated theory of linguistic descriptions.* Cambridge, MA: MIT Press.

Kawamoto, Shigeo (1913–1983)

Y Nagashima, Gakushuin University, Tokyo, Japan

Shigeo Kawamoto was born on January 3, 1913, in Tokyo. He graduated from the Department of English, Faculty of Literature, Waseda University, in 1937. After 3 years in Paris as a *boursier* (1937–1939), where he studied general linguistics and Old French, he returned to Japan and was appointed assistant professor in 1946, professor of French and linguistics from 1949 at the Faculty of Literature, Waseda University, and professor emeritus in 1983. He was founder and first president of the Semiotic Society of Japan (1980–1983), and he served as president of the Linguistic Society of Japan (1981–1983). He died on August 1, 1983.

His major fields of study include French and Japanese linguistics, general linguistics, poetics, and semiotics. His scholarly work began with an analysis of the French verbal system (Kawamoto, 1954, 1955). Later, as a result of a meeting with Roman Jakobson, he took a turn toward generative grammar, poetics, and semiotics (Kawamoto, 1967, 1969, 1981), and he translated Chomsky's *Cartesian linguistics* (1970), *Language and mind* (1975), and *Problem of knowledge and freedom* (1975). He also edited the translation of Jakobson's *Essais de linguistique général* (1973) and *Poetics*, the third volume of the *Selected writings of R. Jakobson* (1978). With his excellent practical knowledge of French and English, he compiled the *Concise French–Japanese dictionary* (1953) and *A new world English–Japanese dictionary* (1969). Three collections of his linguistic essays (Kawamato, 1976, 1978, 1979) show the wide range of his approach to Japanese and language in general.

See also: Chomsky, Noam (b. 1928); Hattori, Shiro (1908–1995); Jakobson, Roman (1896–1982); Japan: History of Linguistics.

Bibliography

Inoue K, Furukawa H & Toyama I (1984). 'Obituaries of S. Kawamoto.' *Gengo Kenkyu 85,* 1–21.
Kawamoto S (1954). 'Les temps composés en français.' *Gengo Kenkyu 26/27,* 78–90.
Kawamoto S (1955). 'L'imparfait français.' *Gengo Kenkyu 28,* 63–74.
Kawamoto S (1967). 'The place of modus in transformational grammar.' *Genko Kenkyu 51,* 52–58.
Kawamoto S (1969). 'The passive in Japanese.' *Cahiers Ferdinand de Saussure 26.*
Kawamoto S (1976). *Language and mind.* Tokyo: Iwanami Shoten.
Kawamoto S (1978). *Colors of language.* Tokyo: Iwanami Shoten.
Kawamoto S (1979). *Thoughts on language.* Tokyo: Kodansha Bunko.
Kawamoto S (1981). Linguistics and semiotics. *Gengo Kenkyu 80,* 1–12.

Kay, Martin (b. 1935)

J Léon, Centre National de la Recherche Scientifique, Paris, France

Martin Kay is a well-known computational linguist. He was responsible for introducing the notions of chart parsing in computational linguistics and of unification in linguistics generally.

He was born in 1935 in Great Britain, and graduated in modern and medieval languages at Trinity College, Cambridge, in 1961. In 1958, he joined the Cambridge Language Research Unit, which was created in 1955 as one of the pioneering centers for machine translation. The works of the CLRU, gathering the philosopher Margaret Masterman, who directed the unit, the biologist R. H. Richens, the linguist M. A. K. Halliday, and computer scientists,

Katz J J (2004). *Sense, reference, and philosophy.* New York: Oxford University Press.
Katz J J & Fodor J A (1963). 'The structure of a semantic theory.' *Language 39,* 170–210.
Katz J J & Langendoen D T (1976). 'Pragmatics and presupposition.' *Language 52,* 1–17.
Katz J J & Postal P M (1964). *An integrated theory of linguistic descriptions.* Cambridge, MA: MIT Press.

Kawamoto, Shigeo (1913–1983)

Y Nagashima, Gakushuin University, Tokyo, Japan

Shigeo Kawamoto was born on January 3, 1913, in Tokyo. He graduated from the Department of English, Faculty of Literature, Waseda University, in 1937. After 3 years in Paris as a *boursier* (1937–1939), where he studied general linguistics and Old French, he returned to Japan and was appointed assistant professor in 1946, professor of French and linguistics from 1949 at the Faculty of Literature, Waseda University, and professor emeritus in 1983. He was founder and first president of the Semiotic Society of Japan (1980–1983), and he served as president of the Linguistic Society of Japan (1981–1983). He died on August 1, 1983.

His major fields of study include French and Japanese linguistics, general linguistics, poetics, and semiotics. His scholarly work began with an analysis of the French verbal system (Kawamoto, 1954, 1955). Later, as a result of a meeting with Roman Jakobson, he took a turn toward generative grammar, poetics, and semiotics (Kawamoto, 1967, 1969, 1981), and he translated Chomsky's *Cartesian linguistics* (1970), *Language and mind* (1975), and *Problem of knowledge and freedom* (1975). He also edited the translation of Jakobson's *Essais de linguistique général* (1973) and *Poetics*, the third volume of the *Selected writings of R. Jakobson* (1978). With his excellent practical knowledge of French and English, he compiled the *Concise French–Japanese dictionary* (1953) and *A new world English–Japanese dictionary* (1969). Three collections of his linguistic essays (Kawamato, 1976, 1978, 1979) show the wide range of his approach to Japanese and language in general.

See also: Chomsky, Noam (b. 1928); Hattori, Shiro (1908–1995); Jakobson, Roman (1896–1982); Japan: History of Linguistics.

Bibliography

Inoue K, Furukawa H & Toyama I (1984). 'Obituaries of S. Kawamoto.' *Gengo Kenkyu 85,* 1–21.
Kawamoto S (1954). 'Les temps composés en français.' *Gengo Kenkyu 26/27,* 78–90.
Kawamoto S (1955). 'L'imparfait français.' *Gengo Kenkyu 28,* 63–74.
Kawamoto S (1967). 'The place of modus in transformational grammar.' *Genko Kenkyu 51,* 52–58.
Kawamoto S (1969). 'The passive in Japanese.' *Cahiers Ferdinand de Saussure 26.*
Kawamoto S (1976). *Language and mind.* Tokyo: Iwanami Shoten.
Kawamoto S (1978). *Colors of language.* Tokyo: Iwanami Shoten.
Kawamoto S (1979). *Thoughts on language.* Tokyo: Kodansha Bunko.
Kawamoto S (1981). Linguistics and semiotics. *Gengo Kenkyu 80,* 1–12.

Kay, Martin (b. 1935)

J Léon, Centre National de la Recherche Scientifique, Paris, France

Martin Kay is a well-known computational linguist. He was responsible for introducing the notions of chart parsing in computational linguistics and of unification in linguistics generally.

He was born in 1935 in Great Britain, and graduated in modern and medieval languages at Trinity College, Cambridge, in 1961. In 1958, he joined the Cambridge Language Research Unit, which was created in 1955 as one of the pioneering centers for machine translation. The works of the CLRU, gathering the philosopher Margaret Masterman, who directed the unit, the biologist R. H. Richens, the linguist M. A. K. Halliday, and computer scientists,

Katz J J (2004). *Sense, reference, and philosophy.* New York: Oxford University Press.
Katz J J & Fodor J A (1963). 'The structure of a semantic theory.' *Language 39,* 170–210.
Katz J J & Langendoen D T (1976). 'Pragmatics and presupposition.' *Language 52,* 1–17.
Katz J J & Postal P M (1964). *An integrated theory of linguistic descriptions.* Cambridge, MA: MIT Press.

Kawamoto, Shigeo (1913–1983)

Y Nagashima, Gakushuin University, Tokyo, Japan

Shigeo Kawamoto was born on January 3, 1913, in Tokyo. He graduated from the Department of English, Faculty of Literature, Waseda University, in 1937. After 3 years in Paris as a *boursier* (1937–1939), where he studied general linguistics and Old French, he returned to Japan and was appointed assistant professor in 1946, professor of French and linguistics from 1949 at the Faculty of Literature, Waseda University, and professor emeritus in 1983. He was founder and first president of the Semiotic Society of Japan (1980–1983), and he served as president of the Linguistic Society of Japan (1981–1983). He died on August 1, 1983.

His major fields of study include French and Japanese linguistics, general linguistics, poetics, and semiotics. His scholarly work began with an analysis of the French verbal system (Kawamoto, 1954, 1955). Later, as a result of a meeting with Roman Jakobson, he took a turn toward generative grammar, poetics, and semiotics (Kawamoto, 1967, 1969, 1981), and he translated Chomsky's *Cartesian linguistics* (1970), *Language and mind* (1975), and *Problem of knowledge and freedom* (1975). He also edited the translation of Jakobson's *Essais de linguistique général* (1973) and *Poetics*, the third volume of the *Selected writings of R. Jakobson* (1978). With his excellent practical knowledge of French and English, he compiled the *Concise French–Japanese dictionary* (1953) and *A new world English–Japanese dictionary* (1969). Three collections of his linguistic essays (Kawamato, 1976, 1978, 1979) show the wide range of his approach to Japanese and language in general.

See also: Chomsky, Noam (b. 1928); Hattori, Shiro (1908–1995); Jakobson, Roman (1896–1982); Japan: History of Linguistics.

Bibliography

Inoue K, Furukawa H & Toyama I (1984). 'Obituaries of S. Kawamoto.' *Gengo Kenkyu 85,* 1–21.
Kawamoto S (1954). 'Les temps composés en français.' *Gengo Kenkyu 26/27,* 78–90.
Kawamoto S (1955). 'L'imparfait français.' *Gengo Kenkyu 28,* 63–74.
Kawamoto S (1967). 'The place of modus in transformational grammar.' *Genko Kenkyu 51,* 52–58.
Kawamoto S (1969). 'The passive in Japanese.' *Cahiers Ferdinand de Saussure 26.*
Kawamoto S (1976). *Language and mind.* Tokyo: Iwanami Shoten.
Kawamoto S (1978). *Colors of language.* Tokyo: Iwanami Shoten.
Kawamoto S (1979). *Thoughts on language.* Tokyo: Kodansha Bunko.
Kawamoto S (1981). Linguistics and semiotics. *Gengo Kenkyu 80,* 1–12.

Kay, Martin (b. 1935)

J Léon, Centre National de la Recherche Scientifique, Paris, France

Martin Kay is a well-known computational linguist. He was responsible for introducing the notions of chart parsing in computational linguistics and of unification in linguistics generally.

He was born in 1935 in Great Britain, and graduated in modern and medieval languages at Trinity College, Cambridge, in 1961. In 1958, he joined the Cambridge Language Research Unit, which was created in 1955 as one of the pioneering centers for machine translation. The works of the CLRU, gathering the philosopher Margaret Masterman, who directed the unit, the biologist R. H. Richens, the linguist M. A. K. Halliday, and computer scientists,

among them Karen Sparck-Jones and Yorick Wilks, focused on a method of machine translation using intermediary language.

Since 1961 Martin Kay has been working in the United States. He graduated in linguistics at the University of California, Berkeley, and joined the Rand Corporation in Santa Monica, another early machine translation center, in 1964, to work on parsing programs, eventually becoming head of Rand's research in linguistics and translation. In 1972 he went to the University of California, Irvine, as Chair of Computer Science, and then in 1974 moved to Xerox Research Center in Palo Alto. In 1985 he became a Professor of Linguistics at Stanford, while retaining his position as a Research Fellow at Xerox, which he kept until 2002. Kay is a past president of the Association for Computational Linguistics and has served for many years as the Chairman of the International Committee on Computational Linguistics. He was awarded an Honorary Doctorate from the University of Gothenburg and he is also an Honorary Professor at the University of the Saarland.

Kay has contributed substantially to several areas of computational linguistics, including machine translation, the state and condition of which he has never ceased to criticize. His more recent contribution to this domain is the development of experimental prototypes of the system Verbmobil that provides simultaneous language translations. His other main achievements have focused on formal languages and formalisms and their use in parsing systems, especially unification grammars. He developed chart parsing and showed that chart parsing algorithms could be carried over to the problem of generating sentences. With Ronald Kaplan, he pioneered finite state morphology, which assumes that both the word formation rules and the morpho-phonological alternation rules can be modeled as finite state machines.

See also: Chart Parsing and Well-Formed Substring Tables; Machine Translation: Overview; Unification: Computational Issues.

Bibliography

Kay M (1982). 'Machine translation.' *American Journal of Computational Linguistics 8(2),* 74–78.

Kay M (1985). 'Parsing in functional unification grammar.' In Dowty D, Karttunen L & Zwicky A (eds.) *Natural language processing.* Cambridge: Cambridge University Press. 251–278.

Kay M (1987). 'Unification and the new grammatism: the linguistic connection.' In *Proceedings of the 1987 Workshop on Theoretical Issues in Natural Language Processing.* Association for Computational Linguistics, Morristown, NJ, USA. 51–57.

Kay M (1996). 'Chart generation.' In *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics* (ACL '96). University of California at Santa Cruz, California, USA, 24–27 June 1996. [On-Line Proceedings.] 200–204.

Kay M (1997). 'The proper place of men and machines in language translation.' *Machine Translation 12,* 3–21.

Kay M & Kaplan R M (1994). 'Regular models of phonological rules systems.' *Computational Linguistics 20(3),* 331–378.

Kay M & Sparck-Jones K (1973). *Linguistics and information science.* New York: Academic Press Inc.

Kay M, Norvig P & Cawron M (1992). *Verbmobil: A translation system for face-to-face dialog.* Chicago: University of Chicago Press.

Kayardild

N Evans, University of Melbourne, Melbourne, Australia

Kayardild is spoken in Queensland, Australia, in the South Wellesley Islands, and belongs to the Tangkic family (non-Pama-Nyungan), which also includes Lardil and Yukulta. Additional and now extinct varieties are Yangkaal and Nguburindi, though the limited materials we have on these show Yangkaal to be a sister dialect of Kayardild and Nguburindi a sister dialect of Yukulta. The Tangkic languages have no close relatives, though they are related, at a distant level, to other Australian languages and share most grammatical similarities with languages along the Roper River well to the west (Evans, 1995).

Speakers of Kayardild were traditionally hunter-gatherers, with a strong emphasis on marine resources, building stone walls around the coasts to catch fish and hunting for turtle and dugong. The traditional population numbered between 120 and 150. Isolated from European contacts until the 1940s, the entire tribe was removed from their homeland in the early 1940s to the mission on Mornington Island, from which date rapid language loss set in: no one born

among them Karen Sparck-Jones and Yorick Wilks, focused on a method of machine translation using intermediary language.

Since 1961 Martin Kay has been working in the United States. He graduated in linguistics at the University of California, Berkeley, and joined the Rand Corporation in Santa Monica, another early machine translation center, in 1964, to work on parsing programs, eventually becoming head of Rand's research in linguistics and translation. In 1972 he went to the University of California, Irvine, as Chair of Computer Science, and then in 1974 moved to Xerox Research Center in Palo Alto. In 1985 he became a Professor of Linguistics at Stanford, while retaining his position as a Research Fellow at Xerox, which he kept until 2002. Kay is a past president of the Association for Computational Linguistics and has served for many years as the Chairman of the International Committee on Computational Linguistics. He was awarded an Honorary Doctorate from the University of Gothenburg and he is also an Honorary Professor at the University of the Saarland.

Kay has contributed substantially to several areas of computational linguistics, including machine translation, the state and condition of which he has never ceased to criticize. His more recent contribution to this domain is the development of experimental prototypes of the system Verbmobil that provides simultaneous language translations. His other main achievements have focused on formal languages and formalisms and their use in parsing systems, especially unification grammars. He developed chart parsing and showed that chart parsing algorithms could be carried over to the problem of generating sentences. With Ronald Kaplan, he pioneered finite state morphology, which assumes that both the word formation rules and the morpho-phonological alternation rules can be modeled as finite state machines.

See also: Chart Parsing and Well-Formed Substring Tables; Machine Translation: Overview; Unification: Computational Issues.

Bibliography

Kay M (1982). 'Machine translation.' *American Journal of Computational Linguistics 8(2),* 74–78.

Kay M (1985). 'Parsing in functional unification grammar.' In Dowty D, Karttunen L & Zwicky A (eds.) *Natural language processing.* Cambridge: Cambridge University Press. 251–278.

Kay M (1987). 'Unification and the new grammatism: the linguistic connection.' In *Proceedings of the 1987 Workshop on Theoretical Issues in Natural Language Processing.* Association for Computational Linguistics, Morristown, NJ, USA. 51–57.

Kay M (1996). 'Chart generation.' In *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics* (ACL '96). University of California at Santa Cruz, California, USA, 24–27 June 1996. [On-Line Proceedings.] 200–204.

Kay M (1997). 'The proper place of men and machines in language translation.' *Machine Translation 12,* 3–21.

Kay M & Kaplan R M (1994). 'Regular models of phonological rules systems.' *Computational Linguistics 20(3),* 331–378.

Kay M & Sparck-Jones K (1973). *Linguistics and information science.* New York: Academic Press Inc.

Kay M, Norvig P & Cawron M (1992). *Verbmobil: A translation system for face-to-face dialog.* Chicago: University of Chicago Press.

Kayardild

N Evans, University of Melbourne, Melbourne, Australia

Kayardild is spoken in Queensland, Australia, in the South Wellesley Islands, and belongs to the Tangkic family (non-Pama-Nyungan), which also includes Lardil and Yukulta. Additional and now extinct varieties are Yangkaal and Nguburindi, though the limited materials we have on these show Yangkaal to be a sister dialect of Kayardild and Nguburindi a sister dialect of Yukulta. The Tangkic languages have no close relatives, though they are related, at a distant level, to other Australian languages and share most grammatical similarities with languages along the Roper River well to the west (Evans, 1995).

Speakers of Kayardild were traditionally hunter-gatherers, with a strong emphasis on marine resources, building stone walls around the coasts to catch fish and hunting for turtle and dugong. The traditional population numbered between 120 and 150. Isolated from European contacts until the 1940s, the entire tribe was removed from their homeland in the early 1940s to the mission on Mornington Island, from which date rapid language loss set in: no one born

after the move to Mornington grew up to be a fluent speaker, and today fewer than 10 speakers remain.

Apart from scanty early word lists, all materials on the Tangkic languages were recorded since the early 1960s. Practical orthographies were developed in this period. These use digraphs for a variety of phonemes, making use of *r* before a stop or nasal letter to denote retroflexion: thus *rd* for /ʈ/ and *rn* for /ɳ/; *h* after a stop or nasal letter to denote a laminointerdental articulation (with the blade of the tongue between the teeth); thus *th* for /t̪/ and *nh* for /n̪/. Other graphemes, standard in Australian orthographies, are *ng* for /ŋ/, *ny* for /ɲ/, *rr* for trilled or flapped /r/, *r* for /l/ and *j* for /c/. Distinctive vowel length is shown by doubling the letter, e.g., *aa* for /a:/.

Phonologically, Kayardild is a typical Australian language, with paired stops and nasals at six points of articulation (bilabial, velar, laminodental, laminopalatal, apicoalveolar and apicoretroflex), a single stop series without voicing contrast, no fricatives, two rhotics (a glide and a tap/trill), and a simple vowel inventory: three vowels (*a*, *i*, *u*) plus length. Primary stress falls on the first syllable unless attracted onto a long vowel.

Kayardild is typical of Australian languages in employing a rich system of case suffixes, which allow for great freedom of word order. Beyond this, the case systems of Kayardild and the other Tangkic languages are remarkable for several reasons. Firstly they exhibit 'double case marking' (see Dench and Evans, 1988; Plank, 1995), since one NP embedded in another inflects both for its own case (e.g., the possessive) and that of the head: cf. Kayardild *thabuju-karra* [brother-POSSESSIVE] 'brother's,' *wangal-nguni* [boomerang-INSTRUMENTAL] 'with the boomerang,' *thabuju-karra-nguni wangal-nguni* 'with brother's boomerang.' Secondly, Kayardild and Lardil add a further 'modal case' inflection, etymologically a case suffix, which marks tense/mood on most nonsubject NPs as a partly parallel system to the tense/mood inflection on the verb. Example (1) gives a Kayardild example using the 'modal ablative' (glossed M.ABL) to mark past tense on the object and instrument NPs in addition to the verbal 'past' inflection.

(1) dangka-a burldi-jarra yarbuth-ina
man-NOM hit-PST bird-M.ABL
thabuju-karra-nguni-na wangal-nguni-na
brother-GEN-INST-M.ABL boomerang-INST-M.ABL
'the man hit the bird with brother's boomerang.'

Thirdly, Kayardild has a further 'complementizing' use of case suffixes, to indicate various types of interclausal relations such as being a clausal complement, as illustrated by the 'complementizing' use of the oblique in (2); note that it goes on all words of the subordinate clause, outside any other inflections. It is also used – on all words except the topicalized object – in strings of topic chains.

(2) ngada kurri-ja, dangka-ntha
1sgNOM see-PST man-NOM
burldi-jarra-ntha yarbuth-ina
hit-PST-C.OBL bird-M.ABL-C.OBL
thabuju-karra-nguni-na wangal-nguni-na
brother-GEN-INST-M.ABL-C.OBL boomerang-INST-M.ABL-C.OBL
'I saw that the man had hit the bird with brother's boomerang.'

Finally, Kayardild can add a layer of 'associating case' on all nonsubject NPs of clauses whose verb has been nominalized.

Much of the how this strange system evolved has now been reconstructed with the help of data from Yukulta, the most conservative Tangkic language and representative of the proto-Tangkic situation in having an ergative: absolutive case system; see Evans (1995) for a summary. Essentially, the main clause structures found in Kayardild and Lardil result from either reanalysis of alternative semitransitive structures in Yukulta, or from the generalization of Yukulta subordinate clause morphology, with case-marking interclausal relations, e.g., proprietive for purposive or ablative for prior time, smeared over both the subordinate verb and its overt NPs. A series of catastrophic changes has thus led Kayardild and Lardil to have grammatical systems that are organized quite differently from that in Yukulta, despite the fact that almost all grammatical morphemes are cognate. These changes make the Tangkic languages a fascinating case of radical and intertwined diachronic developments linked to the abandonment of an ancestral system of ergative case marking.

Another strange feature of Kayardild is a further set of case inflections, semantically and structurally part of the set of 'normal' case inflections but with the peculiarity that they convert their hosts, morphologically, from nouns into verbs. Beneficiaries, for example, take the 'verbal dative' case *-maru-*, which then takes regular verbal inflections (3), but which is distributed across all words in the noun phrase like a case inflection. Etymologically this derives from a verb meaning 'put' but structurally it is now a part of the regular system of case suffixes.

(3) ngada waa-jarra wangarr-ina
1SgNOM sing-PST song-M.ABL
ngijin-maru-tharra thabuju-maru-tharra.
my-V.DAT-PST brother-V.DAT-PST
'I sang a song for my brother.'

Like many other Australian languages, Kayardild has a complex set of derivatives from compass terms. To locate an entity one normally says things like 'the

east uncle' or 'the groper coming from the east'; some examples of derivatives based on the root ri- 'east' are *riinda* 'coming from the east,' *rilungka* 'eastwards,' *riliida* 'heading ever eastwards,' *riyananganda* 'to the east of,' *ringurrnga* 'east across a geographical boundary,' *riyanyinda* 'at the eastern extremity of,' *rilumirdamirda* 'sea-grass territory to the east,' *rilurayaanda* 'from one's previous night's camp in the east,' *rilijulutha* 'move to the east' and *riinmali* 'hey you coming from the east.'

See also: Australia: Language Situation; Australian Languages; Endangered Languages; Future Tense and Future Time Reference; Tense; Word Classes/Parts of Speech: Overview.

Bibliography

Dench A & Evans N D (1988). 'Multiple case-marking in Australian languages.' *Australian Journal of Linguistics 8*, 1–47.
Evans N D (1995). *A grammar of Kayardild, with historical-comparative notes on Tangkic.* Berlin: Mouton de Gruyter.
Plank F (ed.) (1995). *Double case: agreement by Suffixaufnahme.* Oxford: Oxford University Press.

Kaytetye

H Koch, Australian National University, Canberra, Australia

Kaytetye is the only member of the northern branch of the Arandic subgroup of the Pama-Nyungan language family, whose southern branch includes Arrernte (Aranda) and other languages (Hale, 1962; Koch, 2004). It is spoken in the southern part of the Northern Territory of Australia around Barrow Creek and Wauchope. Its speakers number around 200. Most contemporary people of Kaytetye descent speak, in addition to or instead of Kaytetye, one of the neighboring languages – Alyawarr, Anmatyerre (both Arandic), or Warlpiri – and varieties of English. The current language of young people differs in a number of respects from that documented in the 1960s and 1970s by Hale and Koch.

Kaytetye phonology includes the typical Australian consonants, plus a set of pre-stopped nasals, rounded consonants, and a velar glide (the unrounded counterpart to *w*). An even more rare feature is a set of pre-palatalized apical consonants. The system of unrounded consonants and their orthographic representation is shown in **Table 1**. Rounded consonants are indicated by *w* after the consonant (cluster). The vowel system consists of just /a/, a high/mid-central vowel (spelled *e* in the orthography), and a marginal /i/. There is no rounded vowel phoneme, but /e/ has rounded allophones.

Atypical phonotactic features include common word-initial V(C) syllables and final CV syllables with no coda consonant and an obligatory vowel /e/; only word-internal syllables have the full structure CV(C). Word stress falls on the first CV(C) syllable. A word like *arrkwentyarte* 'three' illustrates these features (where the stressed syllable is underlined). The atypical phonology results from a series of sound changes that are described in Koch (1997, 2004).

Nouns inflect for number and case, in the typical Australian fashion, but with some complications (Koch, 1990). Kin nouns, as illustrated in **Table 2**, are inflected for the (singular) person of possessor, by means of prefixes (former dative pronouns) for second and third person, and a suffix or zero for first person. A suffix *-nhenge* marks the category 'dyadic', designating both persons in a relationship, e.g., 'mother and child(ren)'.

Dual and plural personal pronouns, as seen in **Table 3**, mark a kinship-related category called

Table 1 Unrounded consonants (atypical segments in boldface)

	Labial	*Velar*	*Lamino-dental*	*Lamino-palatal*	*Apico-alveolar*	*Apico-postalveolar*	*Pre-palatalized apical*
Stops	p	k	th	ty	t	rt	**yt**
Prestopped nasals	**pm**	**kng**	**tnh**	**tny**	**tn**	**rtn**	**ytn**
Plain nasals	m	ng	nh	ny	n	rn	**yn**
Laterals			lh	ly	l	rl	**yl**
Tap/trill					rr		
Approximants		**h**		y		r	

east uncle' or 'the groper coming from the east'; some examples of derivatives based on the root ri- 'east' are *riinda* 'coming from the east,' *rilungka* 'eastwards,' *riliida* 'heading ever eastwards,' *riyananganda* 'to the east of,' *ringurrnga* 'east across a geographical boundary,' *riyanyinda* 'at the eastern extremity of,' *rilumirdamirda* 'sea-grass territory to the east,' *rilurayaanda* 'from one's previous night's camp in the east,' *rilijulutha* 'move to the east' and *riinmali* 'hey you coming from the east.'

See also: Australia: Language Situation; Australian Languages; Endangered Languages; Future Tense and Future Time Reference; Tense; Word Classes/Parts of Speech: Overview.

Bibliography

Dench A & Evans N D (1988). 'Multiple case-marking in Australian languages.' *Australian Journal of Linguistics 8*, 1–47.
Evans N D (1995). *A grammar of Kayardild, with historical-comparative notes on Tangkic*. Berlin: Mouton de Gruyter.
Plank F (ed.) (1995). *Double case: agreement by Suffixaufnahme*. Oxford: Oxford University Press.

Kaytetye

H Koch, Australian National University, Canberra, Australia

Kaytetye is the only member of the northern branch of the Arandic subgroup of the Pama-Nyungan language family, whose southern branch includes Arrernte (Aranda) and other languages (Hale, 1962; Koch, 2004). It is spoken in the southern part of the Northern Territory of Australia around Barrow Creek and Wauchope. Its speakers number around 200. Most contemporary people of Kaytetye descent speak, in addition to or instead of Kaytetye, one of the neighboring languages – Alyawarr, Anmatyerre (both Arandic), or Warlpiri – and varieties of English. The current language of young people differs in a number of respects from that documented in the 1960s and 1970s by Hale and Koch.

Kaytetye phonology includes the typical Australian consonants, plus a set of pre-stopped nasals, rounded consonants, and a velar glide (the unrounded counterpart to *w*). An even more rare feature is a set of pre-palatalized apical consonants. The system of unrounded consonants and their orthographic representation is shown in Table 1. Rounded consonants are indicated by *w* after the consonant (cluster). The vowel system consists of just /a/, a high/mid-central vowel (spelled *e* in the orthography), and a marginal /i/. There is no rounded vowel phoneme, but /e/ has rounded allophones.

Atypical phonotactic features include common word-initial V(C) syllables and final CV syllables with no coda consonant and an obligatory vowel /e/; only word-internal syllables have the full structure CV(C). Word stress falls on the first CV(C) syllable. A word like *arr<u>kwen</u>tyarte* 'three' illustrates these features (where the stressed syllable is underlined). The atypical phonology results from a series of sound changes that are described in Koch (1997, 2004).

Nouns inflect for number and case, in the typical Australian fashion, but with some complications (Koch, 1990). Kin nouns, as illustrated in Table 2, are inflected for the (singular) person of possessor, by means of prefixes (former dative pronouns) for second and third person, and a suffix or zero for first person. A suffix *-nhenge* marks the category 'dyadic', designating both persons in a relationship, e.g., 'mother and child(ren)'.

Dual and plural personal pronouns, as seen in Table 3, mark a kinship-related category called

Table 1 Unrounded consonants (atypical segments in boldface)

	Labial	*Velar*	*Lamino-dental*	*Lamino-palatal*	*Apico-alveolar*	*Apico-postalveolar*	*Pre-palatalized apical*
Stops	p	k	th	ty	t	rt	**yt**
Prestopped nasals	**pm**	**kng**	**tnh**	**tny**	**tn**	**rtn**	**ytn**
Plain nasals	m	ng	nh	ny	n	rn	**yn**
Laterals			lh	ly	l	rl	**yl**
Tap/trill					rr		
Approximants		**h**		y		r	

Table 2 Kin noun inflection

	'Elder brother'	*'Mother'*
1Sg	alkere-ye	arrwengke_
2Sg	ngk-alkere	ngk-arrwengke
3Sg	kw-alkere	kw-arrwengke
Dyadic	alkere-nhenge	arrwengke-nhenge

Table 3 Twelve words for 'we'

	Same moiety, same generation	*Same moiety, opposite generation*	*Opposite moiety*
Dual inclusive	ayleme	aylake	aylanthe
Dual exclusive	aylene	aylenake	aylenanthe
Plural inclusive	aynangke	aynake	aynanthe
Plural exclusive	aynenangke	aynenake	aynenanthe

Table 4 Associated motion stems for intransitive and transitive verbs

Relative time of motion	*Gloss*	Angke *'talk'*	Kwathe *'drink'*
Prior	VERB after SUBJ goes	angkeyene-	kwatheyene-
	VERB after SUBJ comes	angkeyetnye-	kwatheyetnye-
	VERB after SUBJ returns	angkeyalpe-	kwatheyalpe-
	VERB after non-SUBJ arrives	angkeyayte-	kwatheyayte-
Subsequent	VERB before SUBJ goes away	angkerrayte-	kwathelayte-
	VERB before SUBJ returns	angkerralpe-	kwathelalpe-
Concurrent	VERB while SUBJ comes	angkeyernalpe-	kwatheyernalpe-
	VERB while SUBJ goes along	angkerrape-	kwatherrapeyne-
	VERB continuously/repeatedly while SUBJ goes along	angkerrangkerrenye-	kwathelathelarre-
	VERB once while SUBJ is on the way	angkelpangke-	kwathelpathe-
Prior and subsequent	go and VERB and then return	angkenyayne-	kwathenyayne-

'section' (Koch, 1982), which distinguishes whether the participants belong to different patrimoieties (e.g., *I and my mother/spouse/sister's child*) or the same moiety; and if the latter, to the same or opposite set of alternate generation levels (e.g., *I and my siblings/grandparents/grandchildren* vs. *I and my father/brother's children*).

Verbs may inflect for the category of 'associated motion' (Koch, 1984), indicating distinctions in the direction and relative timing of movement by thesubject (usually), using markers that partially differ according to the transitivity of the verb stem (Table 4) and derive in part from former verbs of motion.

In semantics, the expression of feelings and emotions is characterized by the use of the reflexive construction of 'hear' and the mention of body parts, especially *aleme* 'stomach' (Turpin, 2002).

> *Rewenhe aleme eyterrtye-le elpathe-nke errpatye*
> 3RDSING.REFL stomach person-ERG hear-PRES bad
> 'That person feels bad'

Available language resources include a non-technical learner's guide (Turpin, 2000) and picture dictionary (Turpin and Ross, 2004), each of which includes an audio component, and a text collection (Thompson, 2003)

See also: Arrernte; Australia: Language Situation; Warlpiri.

Bibliography

Hale K (1962). 'Internal relationships in Arandic of Central Australia.' In Capell A (ed.) *Some linguistic types in Australia* (Oceania Linguistic Monographs 7). Sydney: University of Sydney. 171–183.

Koch H (1982). 'Kinship categories in Kaytej pronouns.' In Heath J, Merlan F & Rumsey A (eds.) *Languages of kinship in Aboriginal Australia* (Oceania Linguistic Monographs 24). Sydney: University of Sydney. 64–71.

Koch H (1984). 'The category of 'associated motion' in Kaytej.' *Language in Central Australia 1*, 23–34.

Koch H (1990). 'Do Australian languages really have morphemes? Issues in Kaytej morphology.' In Austin P, Dixon R M W & Dutton T (eds.) *Language and history: essays in honour of Luise A. Hercus* (Pacific Linguistics C-116). Canberra: Australian National University. 193–208.

Koch H (1997). 'Pama-Nyungan reflexes in the Arandic languages.' In Tryon D & Walsh M (eds.) *Boundary rider: essays in honour of Geoffrey O'Grady* (Pacific Linguistics C-136). Canberra: Australian National University. 271–302.

Koch H (2004). 'The Arandic subgroup of Australian languages.' In Bowern C & Koch H (eds.) *Australian languages: classification and the comparative method*. Amsterdam/Philadelphia: John Benjamins. 127–150.

Thompson T K (2003). *Growing up Kaytetye: stories*. Turpin M (comp.). Alice Springs: IAD Press.

Turpin M (2000). *A learner's guide to Kaytetye*. Alice Springs: IAD Press.

Turpin M (2002). 'Body part terms in Kaytetye feeling expressions.' In Enfield N J & Wierzbicka A (eds.) *The body in description of emotion*. Amsterdam/Philadelphia: John Benjamins. 271–305.

Turpin M & Ross A (2004). *Kaytetye picture dictionary*. Alice Springs: IAD Press.

Kazakh

L Johanson, Johannes Gutenberg University, Mainz, Germany

Location and Speakers

Kazakh (*qazaq tili*, *qazaqša*) belongs to the northwestern or Kipchak branch of the Turkic language family, more specifically to its southern or Aralo-Caspian group. Until the early 20th century, it was called Kazak-Kirghiz, whereas Kirghiz was referred to as Kara-Kirghiz. Kazakh is primarily spoken in the Republic of Kazakhstan (*Qazaqstan Respublikasï*), a vast country situated at the center of the West Eurasian steppe zone. It borders on Turkmenistan, Uzbekistan, Kyrgyzstan, and China in the south, and on the Russian Federation in the north and west. Kazakh is also spoken by minorities in Xinjiang (China), Uzbekistan, Mongolia, Turkmenistan, Kyrgyzstan, the Russian Federation, Tajikistan, Afghanistan, etc. The number of speakers is at least 10 million. There are more than seven million in Kazakhstan, more than one million in Xinjiang, and almost one million in Uzbekistan.

Kazakh is, along with Russian, the official language of the Republic of Kazakhstan. Kazakh–Russian bilingualism is widespread. Though Kazakhs constitute half of the population of the republic, many have a low proficiency in their mother tongue. Russians make up 37% of the population. The declaration of Kazakh as the state language in 1989 was met by protests from the non-Kazakh population. In 1995, Russian was proclaimed the language of interethnic communication. Russian has a dominant status in public life as the main language of instruction, science, business, and communication in professional domains.

Karakalpak (*qaraqalpaqša*), an independent language in the political sense, is a slightly Uzbekicized variety of Kazakh. It is spoken by c. 450 000 persons, mainly in the Autonomous Republic of Karakalpakstan (*Qaraqalpaqstan Respublikasi*) in Uzbekistan, on the lower course and in the delta of the Amudarya River. Small groups of speakers live in other regions, e.g., in the Khorezm and Fergana regions of Uzbekistan and in the Dashkhowuz region of Turkmenistan.

Origin and History

Kazakh goes back to the Kipchak varieties of Uzbek tribes, who founded a huge steppe empire in the second part of the 15th century. Separatist Kazakh tribes split off from the Uzbeks during their migrations, and moved into the northern steppe regions. The ancestors of the Karakalpaks belonged to one of the most important confederations under the Golden Horde. At the beginning of the 17th century, the Kazakhs occupied Tashkent, which remained their capital until 1723. In the 18th century, Kazakh and western Mongol tribes fought for supremacy in the steppes between the Altay mountains and the Caspian Sea. The Kazakh empire disintegrated into three so-called hordes, of which the Great Horde submitted to Russia in 1717. The Kazakh territory was incorporated into Russia in the mid-19th century. A Kazakh constituent republic of the Soviet Union was established in 1920. Kazakhstan declared its sovereignty in 1990 and its full independence in 1991.

Related Languages and Language Contacts

Kazakh is closely related to Karakalpak, Kipchak Uzbek and Noghay (Nogai). Kipchak Uzbek, a vanishing variety formerly spoken mainly in the north and northwest of Uzbekistan, goes back to the original language of the Uzbek nomads. Kazakh has had old close contacts with Mongolic languages, and the recent influence on it of Russian has been strong. The southern part of Kazakhstan constitutes an intensive contact zone with Uzbek (Northern Uzbek). The Kazakh varieties spoken in China have been relatively little influenced by Chinese.

Koch H (1997). 'Pama-Nyungan reflexes in the Arandic languages.' In Tryon D & Walsh M (eds.) *Boundary rider: essays in honour of Geoffrey O'Grady* (Pacific Linguistics C-136). Canberra: Australian National University. 271–302.
Koch H (2004). 'The Arandic subgroup of Australian languages.' In Bowern C & Koch H (eds.) *Australian languages: classification and the comparative method*. Amsterdam/Philadelphia: John Benjamins. 127–150.
Thompson T K (2003). *Growing up Kaytetye: stories*. Turpin M (comp.). Alice Springs: IAD Press.
Turpin M (2000). *A learner's guide to Kaytetye*. Alice Springs: IAD Press.
Turpin M (2002). 'Body part terms in Kaytetye feeling expressions.' In Enfield N J & Wierzbicka A (eds.) *The body in description of emotion*. Amsterdam/ Philadelphia: John Benjamins. 271–305.
Turpin M & Ross A (2004). *Kaytetye picture dictionary*. Alice Springs: IAD Press.

Kazakh

L Johanson, Johannes Gutenberg University, Mainz, Germany

Location and Speakers

Kazakh (*qazaq tili*, *qazaqša*) belongs to the northwestern or Kipchak branch of the Turkic language family, more specifically to its southern or Aralo-Caspian group. Until the early 20th century, it was called Kazak-Kirghiz, whereas Kirghiz was referred to as Kara-Kirghiz. Kazakh is primarily spoken in the Republic of Kazakhstan (*Qazaqstan Respublikasï*), a vast country situated at the center of the West Eurasian steppe zone. It borders on Turkmenistan, Uzbekistan, Kyrgyzstan, and China in the south, and on the Russian Federation in the north and west. Kazakh is also spoken by minorities in Xinjiang (China), Uzbekistan, Mongolia, Turkmenistan, Kyrgyzstan, the Russian Federation, Tajikistan, Afghanistan, etc. The number of speakers is at least 10 million. There are more than seven million in Kazakhstan, more than one million in Xinjiang, and almost one million in Uzbekistan.

Kazakh is, along with Russian, the official language of the Republic of Kazakhstan. Kazakh–Russian bilingualism is widespread. Though Kazakhs constitute half of the population of the republic, many have a low proficiency in their mother tongue. Russians make up 37% of the population. The declaration of Kazakh as the state language in 1989 was met by protests from the non-Kazakh population. In 1995, Russian was proclaimed the language of interethnic communication. Russian has a dominant status in public life as the main language of instruction, science, business, and communication in professional domains.

Karakalpak (*qaraqalpaqša*), an independent language in the political sense, is a slightly Uzbekicized variety of Kazakh. It is spoken by c. 450 000 persons, mainly in the Autonomous Republic of Karakalpakstan (*Qaraqalpaqstan Respublikasi*) in Uzbekistan, on the lower course and in the delta of the Amudarya River. Small groups of speakers live in other regions, e.g., in the Khorezm and Fergana regions of Uzbekistan and in the Dashkhowuz region of Turkmenistan.

Origin and History

Kazakh goes back to the Kipchak varieties of Uzbek tribes, who founded a huge steppe empire in the second part of the 15th century. Separatist Kazakh tribes split off from the Uzbeks during their migrations, and moved into the northern steppe regions. The ancestors of the Karakalpaks belonged to one of the most important confederations under the Golden Horde. At the beginning of the 17th century, the Kazakhs occupied Tashkent, which remained their capital until 1723. In the 18th century, Kazakh and western Mongol tribes fought for supremacy in the steppes between the Altay mountains and the Caspian Sea. The Kazakh empire disintegrated into three so-called hordes, of which the Great Horde submitted to Russia in 1717. The Kazakh territory was incorporated into Russia in the mid-19th century. A Kazakh constituent republic of the Soviet Union was established in 1920. Kazakhstan declared its sovereignty in 1990 and its full independence in 1991.

Related Languages and Language Contacts

Kazakh is closely related to Karakalpak, Kipchak Uzbek and Noghay (Nogai). Kipchak Uzbek, a vanishing variety formerly spoken mainly in the north and northwest of Uzbekistan, goes back to the original language of the Uzbek nomads. Kazakh has had old close contacts with Mongolic languages, and the recent influence on it of Russian has been strong. The southern part of Kazakhstan constitutes an intensive contact zone with Uzbek (Northern Uzbek). The Kazakh varieties spoken in China have been relatively little influenced by Chinese.

The Written Language

Kazakh did not possess a written variety in the pre-Russian period. Official documents were mostly written in Chaghatay (Chagatai) or Tatar. At the end of the 19th century a written language emerged. It was based on the dialect of the northwestern regions, where the Russian and Tatar influence was strong.

The Arabic script was used up to 1929. It was replaced by a Roman-based alphabet, which was abandoned in 1940 in favor of a Cyrillic alphabet. There are currently plans to adopt a Roman-based script again. In China, Kazakh is written in Arabic script, after an unsuccessful experiment with a Roman-based (pinyin) alphabet in the 1970s. Written Kazakh of China is still oriented towards the standard language used in Kazakhstan.

Karakalpak was established in 1925 as a language in the political sense and as a written language. Its orthography differs considerably from that of Kazakh. After the Arabic and Roman scripts had been employed for a few years, the Cyrillic alphabet was introduced in 1940. The transition to a new Roman-based script has begun.

Distinctive Features

Kazakh exhibits most linguistic features typical of the Turkic family (*see* **Turkic Languages**). It is an agglutinative language with suffixing morphology, sound harmony, and a head-final constituent order. In the following, only a few distinctive features will be dealt with. In the notation of suffixes, capital letters indicate phonetic variation, e.g., *A* = *a*/*e*. Hyphens are used here to indicate morpheme boundaries.

Phonology

Kazakh has a front vowel *æ*, which is lower than the mid-high *e* and restricted to the first syllable. It has emerged through fronting of *a*, e.g., *bæri* 'all.' High vowels are often reduced, relatively short, but not lowered as in Tatar and Bashkir. Initial *e*-, *o*- and *ö*- are often pronounced with a prothetic glide, e.g., *ʸeki* 'two,' *ᵘon* 'ten.' Initial *t*- is mostly preserved, e.g., *til* 'language' (cf. Turkish *dil*). As in most Kipchak languages, final -ɣ is labialized in monosyllabic stems, e.g., *taw* 'mountain' ← *ta:*ɣ. The affricate *č* has developed into the fricative *š*, e.g., *üč* 'three' ← *üš*, whereas *š* has developed into *s*, e.g., *tas* 'stone' ← *taš*. Word-initial *ž*- corresponds to *y*- or *ǰ*- in other Turkic languages, e.g., *žol* 'way' (Turkish *yol*). However, *ǰ*- is found in older Kazakh texts.

According to front vs. back sound harmony, a suffix vowel is back if the preceding syllable has a back vowel, and front if the preceding syllable has a front vowel. This type of sound harmony also affects consonants, e.g., *qar-*ɣ*a* [snow-DAT] 'into the snow,' *köl-ge* [lake-DAT] 'into the lake.'

According to rounded vs. unrounded harmony (labial harmony), a suffix vowel is rounded if the preceding syllable contains a rounded vowel, and unrounded if the preceding syllable contains an unrounded vowel. Spoken Kazakh displays this kind of harmony not only in high suffix vowels, but also in low vowels, e.g., *üy-dö* [house-LOC] 'in the house,' *tüs-kön* [fall-PART] 'fallen,' *öl-gön* [die-PART] 'dead.' However, *o* is not admitted, e.g., *qol-da* [hand-LOC] 'in the hand' instead of **qol-do*. The rounding effect decreases with the distance from the first syllable, e.g., *üy-dö* [house-LOC] 'in the house,' *üy-ümüz-de* [house-POSS.1PL-LOC] 'in our house.' The rounded vs. unrounded harmony is not reflected in the orthography, which does not represent the suffix vowels *ö*, *ü*, and *u*.

Kazakh exhibits numerous consonant changes, mostly assimilations, in clusters containing dentals, liquids and nasals. Suffix-initial *d*-, *l*-, *n*-, and *m*- occur after stem-final vowels and often after nasals, sonorants, and glides, but they are otherwise assimilated. Suffix-initial *d*-, *l*-, and *n*- are assimilated to *t*- after voiceless stem-final consonants, e.g., *at-tar* [horse-PL] 'horses' ← *at-lar*. Suffix-initial *d*- is assimilated to *n*- after stem-final nasals, e.g., *adam-nan* [man-ABL] 'from the man.' Suffix-initial *l*- and *n*- are changed to *d*- after most consonants, e.g., *köz-der* [eye-PL] 'eyes,' *qïz-dïᵑ* [girl-GEN] 'of the girl.' Suffix-initial *m*- is changed to *b*- after stem-final -*z* and -*ž*, e.g., *žaz-ba-* [write-NEG] 'not to write.'

Loanwords are mostly pronounced according to indigenous phonotactic rules. Consonant clusters are dissolved through consonant deletion, or insertion of high prothetic or epenthetic vowels, e.g., *xalïq* 'people' (Arabic *xalq*), *ïras* 'true' (Persian *rast*), *pïr-atsent* 'percent' (Russian *procent*), *telgirep* 'telegraph' (Russian *telegraf*), *kerewet* 'bed' (Russian *krovat'*). The consonants *f* and *x* in loanwords are replaced by *p* and *q*, e.g., *aptobus* 'bus,' *qat* 'letter.' The phonological adaptation was largely reflected in the old Arabic and Roman orthographies, but since the introduction of the Cyrillic script, Russian loanwords are written in their original form.

Grammar

The comparative suffix is -*rAK* (after consonant-final stems -*(I)rAK*), e.g., *ülken-irek* [big-COMP] 'bigger.' There is an instrumental case in -*Men*. The reflexive pronoun *öz* is used attributively or as a noun with possessive suffixes, e.g., *öz üy-im* [self house-POSS.1SG] 'my own house,' *öz-im bar-dïm* [self-POSS.1SG go-PAST-1SG] 'I went myself.'

The old flexion of pronouns is preserved, and not replaced by flexion of the nouns as in some other

Turkic languages. Pronouns ending in *-l*, e.g., *bul* 'this,' replace *-l* with *-n* in most oblique cases. The initial *b-* in *bul* is changed to *m-* in the genitive, accusative, and locative. The demonstrative pronouns *bul*, *osï*, and *mïna* are used for referents within the range of view, while *ol*, *sol*, and *ana* are used for referents out of sight. Some demonstratives exhibit emphatic forms, e.g., *mïna-w*.

The second-person plural marker *sIŋ-dAr* combines the second-person singular marker with the plural suffix. The marker *-sIz* expresses politeness when used to one addressee, whereas *-sIz-dAr* is the corresponding polite plural marker.

The suffixes *-LA-p* and *-LA-GAn*, whose first element *-LA* derives verbs stems from nouns, form approximative and multiplicative numerals, e.g., *žüz-de-gen* 'hundreds.'

Kazakh has a present tense in *-A* plus person markers, e.g., *kel-e-di* [come-PRES-3.SG] 'comes.' It has emerged from a construction containing a converb ending in *-A* + *tur(ur)* 'stands'. The old present tense form in *-(A)r* is mostly used with modal meanings, e.g., *kör-er-miz* [see-AOR-1.PL] 'we will see.' A more focal present tense, i.e. with a narrower focus on the ongoing event, is formed with the converb suffix *-(I)p* and one of the verbs *otïr* 'sits,' *žür* 'goes,' *žatïr* 'lies,' *tur* 'stands,' e.g., *žaz-ïp žatïr* [write-CONV lies] 'is writing.' The suffix *-AtIn* (*-ytIn* after vowel-final stems) forms a habitual past. There is an intentional in *-MAK*, e.g., *kel-mek-piz* [come-INTENT-1PL] 'we want to come,' and a necessitative in *-MAK-šI*, e.g., *sat-paqšï-mïn* [sell-NEC-1SG] 'I must sell.' Kazakh displays complex verbal compositions expressing actionality, aspect-tense, and evidentiality. A number of auxiliary verbs are used in postverb constructions based on the converbs in *-A* and *-(I)p* and mostly specifying the manner of action, e.g., *žan-ïp ket-* [burn-CONV go] 'to burn down.' Evidential forms are the indirective past in *-(I)p*, with *-DI* in the third person, e.g., *kel-ip-ti* [come-CONV-3SG] 'apparently came.' The evidential (indirective) copula particle *eken* combines with various items, e.g., *kel-e-di eken* [come-PRES-3SG INDIR. COP] 'apparently comes,' *kel-gen eken* [come-PART EV] 'has apparently arrived.' Unlike more Iranicized languages such as Uzbek and Turkmen, Kazakh has a weakly developed system of conjunctions.

Lexicon

The basic vocabulary related to the traditional way of life is of Kipchak Turkic origin. There is also a rich modern Kazakh vocabulary with numerous neologisms, e.g., *xalïqaralïq* 'international.' Many words are of Persian (Farsi) and Arabic origin, introduced via Tatar and Chaghatay, e.g., *bazar* 'market,' *apta* 'week,' *nan* 'bread,' *aqïl* 'intellect,' *γïlïm* 'science,' *maγïna* 'meaning,' *waqït* 'time.' Due to close contacts with other nomadic groups, Kazakh also exhibits many words of Mongolic origin, e.g., *olža* 'booty.' Kazakh of China has a fairly strong early layer of Mongolic loanwords, but relatively few loans from Chinese, at least in the written language. Russian words have been borrowed from the second half of the 19th century on. In the Soviet period, calques on Russian models and neologisms were generally preferred to loanwords. There is now a certain tendency in the written language to reduce the number of Russian loans in favor of words of native or Arabic–Persian origin.

Dialects

In spite of its huge extent, the Kazakh-speaking area exhibits little dialectal variation because of a high degree of mobility of the speakers of Kazakh throughout their history. The standard language is based on the northwestern dialect. The southern and western dialects differ from it in some respects. For example, initial *y-* is often found, e.g., *yaq* 'side' instead of *žaq*. Common Turkic *č* is often found instead of standard Kazakh *š*. Changes of suffix-initial consonants are less common. They are, on the other hand, stronger in the easternmost dialects spoken in China and Mongolia, e.g., *bala-dar* [child-PL] 'children' instead of *bala-lar* [child-PL]. The rural Kipchak Uzbek dialects in the north and northwest of Uzbekistan have been largely de-Kipchakicized and are now practically extinct due to the abandonment of the nomadic lifestyle.

The Karakalpak language displays most of the phonetic and morphophonemic characteristics of Kazakh. Word-initial *ǰ-* is, however, found instead of *ž-*, e.g., *ǰol* 'road.' Suffix-initial *l-* is not assimilated, e.g., *tas-lar* [stone-PL] 'stones' instead of *tas-tar*, *qïz-lar* [girl-PL] 'girls' instead of *qïz-dar*. There are a few morphological differences, stemming from the influence of Oghuz and Uzbek. The future suffix *-AžAK* is a loan from Oghuz. The first- and second-person plural personal markers *-MAn* and *-sAN* instead of Kazakh *-MIn* and *-sIN* are similar to Uzbek *-man* and *-san*. The vocabulary differs to a certain degree from that of Kazakh and contains more words of Arabic and Persian origin. Karakalpak has two main dialects, one northeastern and one southwestern. Certain nominally Karakalpak and Kazakh dialects in the region of Khorezm belong to the Southwestern or Oghuz branch of Turkic.

See also: Altaic Languages; Kazakhstan: Language Situation; Mongolic Languages; Turkic Languages; Uzbek.

Bibliography

Balakaev M B *et al.* (eds.) (1962). *Sorremennyj kazaxskij jazyk*. Alma-Ata: Izdatel'stvo Akademii Nauk Kazaskoj SSR.

Johanson L (2001). 'Kazakh.' In Garry J & Rubino C (eds.) *Facts about the world's languages: an encyclopedia of the world's major languages, past and present*. New York: H. W. Wilson. 376–378.
Kirchner M (1992). *Phonologie des Kasachischen: Untersuchungen anhand von Sprachaufnahmen aus der kasachischen Exilgruppe in Istanbul*. Wiesbaden: Otto Harrassowitz.
Kirchner M (1998). 'Kazakh and Karakalpak.' In Johanson L & Csató É Á (eds.) *The Turkic languages*. London/New York: Routledge. 318–332.
Krueger J R (1980). *Introduction to Kazakh*. Bloomington: Indiana University Press.
Menges K H (1959). 'Die aralo-kaspische Grüppe.' In Deny J *et al.* (eds.) *Philologiae turcicae fundamenta 1*. Wiesbaden: Steiner. 434–488.
Mïrzabekova Q *et al.* (eds.) (1992). *Qazaqša-nemisše sözdik/Kasachisch-deutsches Wörterbuch*. Almaty: Rawan.
Shnitnikov B N (1966). *Kazakh-English dictionary*. London/The Hague/Paris: Mouton.
Wurm S (1951). 'The Karakalpak language.' *Anthropos 46*, 487–610.

Kazakhstan: Language Situation

E Suleimenova, Kazakh National University, Almaty, Kazakhstan
J Smagulova, Kazakh Institute of Management, Economics, and Strategic Research, Almaty, Kazakhstan

The Republic of Kazakhstan became independent in 1991. In 2004, the population of Kazakhstan (2 717 300 sq km) was 14 953 126: 57.19% Kazakhs; 27.24% Russians; 15.57% others; 53.8% of the population speak two or more languages, 46.2% are monolinguals (**Table 1**).

The major religions are Islam and Orthodox Christianity.

Kazakh is the state language. Both Kazakh and Russian can be officially used in state and local government institutions, including in the courts and in the military; in legal, business, customs, and other

Table 1 Kazakhstan. Ethnic groups and knowledge of languages[a]

Ethnic group	*Population size*	*As percent of total population*	*Percent of those who claim to speak*				
			Their native language	*The nonnative language*			
				Kazakh	*Russian*	*English*	*German*
Total	14 953 126	100.00	90.0	11.2	54.8	0.7	0.3
Kazakh	7 985 039	53.40	99.4	–	75.0	0.7	0.2
Russian	4 479 620	29.96	100.0	14.9	–	0.9	0.4
Ukrainian	547 054	3.66	16.1	12.6	99.5	0.5	0.3
Uzbek	370 663	2.48	97.0	80.0	59.2	0.1	0.0
German	353 441	2.36	21.8	15.4	99.3	0.3	–
Tatar	248 954	1.66	37.1	63.6	96.9	0.8	0.2
Uighur	210 365	1.41	81.3	80.5	76.1	0.5	0.1
Belarusian	111 927	0.74	13.5	9.9	99.4	0.4	0.3
Korean	99 665	0.67	25.8	28.8	97.7	1.8	0.3
Turks	75 933	0.55	75.5	74.7	75.9	0.4	0.1
Azeri	78 295	0.52	50.1	63.4	86.0	0.3	0.1
Polish	47 297	0.32	9.1	9.9	98.9	0.5	0.6
Dungan	36 945	0.25	48.9	35.9	95.1	0.3	0.0
Kurdish	32 764	0.22	68.8	61.2	76.9	0.1	0.1
Chechen	31 799	0.21	58.6	39.5	94.1	0.4	0.2
Tajik	25 659	0.17	56.8	67.9	35.8	0.2	0.0
Bashkir	23 225	0.16	28.2	58.3	97.9	0.4	0.2
Moldovan	19 460	0.13	21.8	17.8	99.2	0.3	0.2
Ingush	16 900	0.11	58.2	22.6	96.9	0.6	0.1
Armenian	14 758	0.09	45.4	21.8	97.5	1.4	0.3
Greek	12 703	0.08	26.8	26.4	98.2	1.2	0.3
Chuvash	11 851	0.08	23.7	19.0	99.3	0.4	0.3
Kyrgyz	10 897	0.07	38.3	88.8	80.7	0.7	0.2

[a]The table lists ethnic groups that are 10 000 people and larger. In Kazakhstan there are more than 120 ethnic groups.
Source: Smailov (2000).

Johanson L (2001). 'Kazakh.' In Garry J & Rubino C (eds.) *Facts about the world's languages: an encyclopedia of the world's major languages, past and present*. New York: H. W. Wilson. 376–378.
Kirchner M (1992). *Phonologie des Kasachischen: Untersuchungen anhand von Sprachaufnahmen aus der kasachischen Exilgruppe in Istanbul*. Wiesbaden: Otto Harrassowitz.
Kirchner M (1998). 'Kazakh and Karakalpak.' In Johanson L & Csató É Á (eds.) *The Turkic languages*. London/New York: Routledge. 318–332.
Krueger J R (1980). *Introduction to Kazakh*. Bloomington: Indiana University Press.
Menges K H (1959). 'Die aralo-kaspische Grüppe.' In Deny J *et al.* (eds.) *Philologiae turcicae fundamenta 1*. Wiesbaden: Steiner. 434–488.
Mïrzabekova Q *et al.* (eds.) (1992). *Qazaqša-nemisše sözdik/Kasachisch-deutsches Wörterbuch*. Almaty: Rawan.
Shnitnikov B N (1966). *Kazakh-English dictionary*. London/The Hague/Paris: Mouton.
Wurm S (1951). 'The Karakalpak language.' *Anthropos 46*, 487–610.

Kazakhstan: Language Situation

E Suleimenova, Kazakh National University, Almaty, Kazakhstan
J Smagulova, Kazakh Institute of Management, Economics, and Strategic Research, Almaty, Kazakhstan

The Republic of Kazakhstan became independent in 1991. In 2004, the population of Kazakhstan (2 717 300 sq km) was 14 953 126: 57.19% Kazakhs; 27.24% Russians; 15.57% others; 53.8% of the population speak two or more languages, 46.2% are monolinguals (Table 1).

The major religions are Islam and Orthodox Christianity.

Kazakh is the state language. Both Kazakh and Russian can be officially used in state and local government institutions, including in the courts and in the military; in legal, business, customs, and other

Table 1 Kazakhstan. Ethnic groups and knowledge of languages[a]

Ethnic group	*Population size*	*As percent of total population*	*Percent of those who claim to speak*				
			Their native language	*The nonnative language*			
				Kazakh	*Russian*	*English*	*German*
Total	14 953 126	100.00	90.0	11.2	54.8	0.7	0.3
Kazakh	7 985 039	53.40	99.4	–	75.0	0.7	0.2
Russian	4 479 620	29.96	100.0	14.9	–	0.9	0.4
Ukrainian	547 054	3.66	16.1	12.6	99.5	0.5	0.3
Uzbek	370 663	2.48	97.0	80.0	59.2	0.1	0.0
German	353 441	2.36	21.8	15.4	99.3	0.3	–
Tatar	248 954	1.66	37.1	63.6	96.9	0.8	0.2
Uighur	210 365	1.41	81.3	80.5	76.1	0.5	0.1
Belarusian	111 927	0.74	13.5	9.9	99.4	0.4	0.3
Korean	99 665	0.67	25.8	28.8	97.7	1.8	0.3
Turks	75 933	0.55	75.5	74.7	75.9	0.4	0.1
Azeri	78 295	0.52	50.1	63.4	86.0	0.3	0.1
Polish	47 297	0.32	9.1	9.9	98.9	0.5	0.6
Dungan	36 945	0.25	48.9	35.9	95.1	0.3	0.0
Kurdish	32 764	0.22	68.8	61.2	76.9	0.1	0.1
Chechen	31 799	0.21	58.6	39.5	94.1	0.4	0.2
Tajik	25 659	0.17	56.8	67.9	35.8	0.2	0.0
Bashkir	23 225	0.16	28.2	58.3	97.9	0.4	0.2
Moldovan	19 460	0.13	21.8	17.8	99.2	0.3	0.2
Ingush	16 900	0.11	58.2	22.6	96.9	0.6	0.1
Armenian	14 758	0.09	45.4	21.8	97.5	1.4	0.3
Greek	12 703	0.08	26.8	26.4	98.2	1.2	0.3
Chuvash	11 851	0.08	23.7	19.0	99.3	0.4	0.3
Kyrgyz	10 897	0.07	38.3	88.8	80.7	0.7	0.2

[a]The table lists ethnic groups that are 10 000 people and larger. In Kazakhstan there are more than 120 ethnic groups.
Source: Smailov (2000).



Figure 1 Kazakh speakers in Kazakhstan by oblast (region). Source: Smailov (2000). 'Itogi perepisi naselenia 1999 goda v Respublike Kazakhstan 2: Naselenie Respubliki Kazakhstan po nationalnostam i vladeniu jazykami (The Results of 1999 census in the Republic of Kazakhstan 2: Population the Republic of Kazakhstan by ethnicity and languages).' Almaty: Statistics Agency of the Republic of Kazakhstan.

Table 2 Use of Kazakh and Russian

Domains	*Kazakh*	*Russian*
Government communications	25.0	75.0
Preschools (as percent of total number of preschools)	25.0	75.0
Schools (as percent of total number of students)	49.4	47.2
Universities (as percent of total number of students)	32.0	68.0
Dissertations defended (as percent of total number of dissertations)	14.0	86.0
Newspapers (as percent of single circulation)	18.1	81.9
Magazines (as percent of annual circulation)	24.2	75.8
Printed matter (as percent of total number of copies)	33.3	66.7

Sources: Nysanbayeva (2000), Gosudarstvennaya programma funktsionirovaniya i razvitiya jazykov na 2001–2010 gody (2001), Smailov (2002a,b).

documentation; in road and other signs; in mass media (broadcasting in Kazakh must be no less than half of all broadcasting time). Kazakh and Russian are required school subjects.

The knowledge of Russian is almost universal among non-Russians. Most language minorities underwent language shift and were linguistically Russified. A high language maintenance rate is observed only among Turkic-speaking non-Kazakh groups (they also have a good command of Kazakh). A total of 3.3% of schoolchildren attend minority language schools: Uzbek, Uyghur, Tajik, German, and Tatar. There are radio programs and newspapers in German, Ukrainian, Korean, and Uyghur.

Twenty-five percent of Kazakhs are monolingual Kazakh speakers; 99.1% of Kazakhs claim to know Kazakh, which is viewed as a primary symbol of Kazakh identity; there is a wide discrepancy between actual and stated native language proficiency. The level of knowledge of Kazakh varies from region to region (Figure 1).

State language policy is aimed at the restoration of the functional domains of the Kazakh language, at the acquisition of Kazakh by the general population, and at corpus planning. Support for Kazakh at the state level has decelerated language shift among Kazakhs and has strengthened the position of Kazakh as the language of the government, mass media, education, and science. Russian remains the dominant language (Table 2).

See also: Bilingualism; German; Identity and Language; Kazakh; Korean; Language Maintenance and Shift;

Minorities and Language; Russian; Tajik Persian; Tatar; Uyghur; Uzbek.

Bibliography

The Constitution of the Republic of Kazakhstan (1995). http://cofinder.richmond.edu.

Dave B & Sinott P (2002). 'Demographics and language politics in the 1999 Kazakhstan Census.' *The National Council for Eurasian and European Research Working Paper.* Washington, DC: The National Council for Eurasian and European Research.

Fierman W (1998). 'Language and identity in Kazakhstan: formulations in policy documents 1987–1997.' *Communist and Post-Soviet Studies 31(2),* 171–186.

Gosudarstvennaya programma funktsionirovaniya I razvitiya jazykov na 2001–2010 gody (State program of functioning and developing of languages in 2001–2010) (2001). *Osnovnye Zakonodatel'nye Akty o Jazykah v Respublike Kazakhstan (Main legislative documents on languages in the Republic of Kazakhstan).* Almaty: Jurist.

Nysanbayeva A (2000). Razvitie jazykovoj politiki v Kazakhstane (The development of language policy in Kazakhstan). http://www/kisi/Parts/books.

Olcott M B (1995). *The Kazakhs.* Stanford: Hoover Institution Press.

Osnovnye zakonodatelnye akty o jazykah v Respublike Kazakhstan (The principal legal documents on languages in the Republic of Kazakhstan). Almaty: Jurist.

Suleimenova E & Akanova D (2002). 'Kazakh language and the new language policy in the 1990s.' In *The Turks 6.* Ankara: Yeni Turkiye Publications. 182–189.

Suleimenova E (ed.) (1997). *Jazykovaia politika v Kazakhstane.* (1921–1990). Sbornik dokumentov (Language policy in Kazakhstan (1921–1990). Collection of documents). Almaty: Kazakh University.

Smailov A (ed.) (2000). *Itogi perepisi naselenia 1999 goda v Respublike Kazakhstan. 1, 2.* (The results of 1999 census in the Republic of Kazakhstan, 1, 2). Almaty: Statistics Agency of the Republic of Kazakhstan.

Smailov A (ed.) (2002a). *Kazakhstan: 1991–2002.* Almaty: Statistics Agency of the Republic of Kazakhstan.

Smailov A (ed.) (2002b). *Statistical Yearbook 2002.* Almaty: Statistics Agency of the Republic of Kazakhstan.

Kazama, Kiyozo (b. 1928)

T Maruyama, Nanzan University, Nagoya, Japan

Kiyozo Kazama is a Japanese Indo-Europeanist who succeeded Harushige Kodzu (*see* **Kodzu, Harushige (1908–1973)**) at the University of Tokyo in 1972. Born in traditional downtown Tokyo in 1928, he studied Indo-European comparative linguistics with Prof. Kodzu. After working as a research assistant in the department of linguistics at the University of Tokyo, he taught linguistics at Nagoya University. In 1972 he started to teach Indo-European comparative linguistics at the University of Tokyo, from where he retired with the title of Emeritus Professor in 1989. Meanwhile he studied at the University of Vienna (Universität Wien), funded by a scholarship from the Japanese Ministry of Education (1973–1974).

One of his major fields is Indo-European kinship terms, on which he wrote his doctoral dissertation in 1978. Its revised version is published in Japan as *Studies in Indo-European kinship terms* (1984). This compilation and interpretation of new materials presents a major new synthetic viewpoint of Indo-European kinship terms, going beyond Delbrück's memorial work *Die indogermanischen Verwandtschaftsname*n (1889).

Kazama is the author of a short history of linguistics entitled *The birth of linguistics* (1978), which deals mainly with the linguistic trends from Sir William Jones to the Neogrammarians, and linguists such as Rasmus Rask, Franz Bopp, Jacob Grimm, August Schleicher, Georg Curtius, Karl Brugmann, Hermann Paul, Hermann Osthoff, Johannes Schmidt, Hugo Schuchardt, and Ferdinand de Saussure; their standpoints or methodologies are described in it clearly and succinctly. His more recent work, *In search for Indo-European Urheimat* (1993) is another exciting history pursuing the concept of an Indo-European homeland and its life and culture. He is also an author of two books concerning the Indo-European culture viewed from its vocabulary, namely, *Indo-European culture viewed from the aspect of words related to daily life* (1987) and *Indo-European culture from the aspect of words related to body parts and functioning* (1990). They belong in the tradition of academic works by Otto Schrader, Carl Darling Buck, and Emile Benveniste. In 1977 he had already translated into Japanese *Die Indogermanen* by Otto Schrader and Hans Krahe. His Latin-Greek 'contrastive' grammar, *A primary guide to Latin and Greek* (1998) is also a very useful handbook for younger scholars and students of linguistics.

Gengo Kenkyu, the journal of the Linguistic Society of Japan, has published his various linguistic

Minorities and Language; Russian; Tajik Persian; Tatar; Uyghur; Uzbek.

Bibliography

The Constitution of the Republic of Kazakhstan (1995). http://cofinder.richmond.edu.

Dave B & Sinott P (2002). 'Demographics and language politics in the 1999 Kazakhstan Census.' *The National Council for Eurasian and European Research Working Paper.* Washington, DC: The National Council for Eurasian and European Research.

Fierman W (1998). 'Language and identity in Kazakhstan: formulations in policy documents 1987–1997.' *Communist and Post-Soviet Studies 31(2),* 171–186.

Gosudarstvennaya programma funktsionirovaniya I razvitiya jazykov na 2001–2010 gody (State program of functioning and developing of languages in 2001–2010) (2001). *Osnovnye Zakonodatel'nye Akty o Jazykah v Respublike Kazakhstan (Main legislative documents on languages in the Republic of Kazakhstan).* Almaty: Jurist.

Nysanbayeva A (2000). Razvitie jazykovoj politiki v Kazakhstane (The development of language policy in Kazakhstan). http://www/kisi/Parts/books.

Olcott M B (1995). *The Kazakhs.* Stanford: Hoover Institution Press.

Osnovnye zakonodatelnye akty o jazykah v Respublike Kazakhstan (The principal legal documents on languages in the Republic of Kazakhstan). Almaty: Jurist.

Suleimenova E & Akanova D (2002). 'Kazakh language and the new language policy in the 1990s.' In *The Turks 6.* Ankara: Yeni Turkiye Publications. 182–189.

Suleimenova E (ed.) (1997). *Jazykovaia politika v Kazakhstane.* (1921–1990). Sbornik dokumentov (Language policy in Kazakhstan (1921–1990). Collection of documents). Almaty: Kazakh University.

Smailov A (ed.) (2000). *Itogi perepisi naselenia 1999 goda v Respublike Kazakhstan. 1, 2.* (The results of 1999 census in the Republic of Kazakhstan, 1, 2). Almaty: Statistics Agency of the Republic of Kazakhstan.

Smailov A (ed.) (2002a). *Kazakhstan: 1991–2002.* Almaty: Statistics Agency of the Republic of Kazakhstan.

Smailov A (ed.) (2002b). *Statistical Yearbook 2002.* Almaty: Statistics Agency of the Republic of Kazakhstan.

Kazama, Kiyozo (b. 1928)

T Maruyama, Nanzan University, Nagoya, Japan

Kiyozo Kazama is a Japanese Indo-Europeanist who succeeded Harushige Kodzu (*see* **Kodzu, Harushige (1908–1973)**) at the University of Tokyo in 1972. Born in traditional downtown Tokyo in 1928, he studied Indo-European comparative linguistics with Prof. Kodzu. After working as a research assistant in the department of linguistics at the University of Tokyo, he taught linguistics at Nagoya University. In 1972 he started to teach Indo-European comparative linguistics at the University of Tokyo, from where he retired with the title of Emeritus Professor in 1989. Meanwhile he studied at the University of Vienna (Universität Wien), funded by a scholarship from the Japanese Ministry of Education (1973–1974).

One of his major fields is Indo-European kinship terms, on which he wrote his doctoral dissertation in 1978. Its revised version is published in Japan as *Studies in Indo-European kinship terms* (1984). This compilation and interpretation of new materials presents a major new synthetic viewpoint of Indo-European kinship terms, going beyond Delbrück's memorial work *Die indogermanischen Verwandtschaftsnamen* (1889).

Kazama is the author of a short history of linguistics entitled *The birth of linguistics* (1978), which deals mainly with the linguistic trends from Sir William Jones to the Neogrammarians, and linguists such as Rasmus Rask, Franz Bopp, Jacob Grimm, August Schleicher, Georg Curtius, Karl Brugmann, Hermann Paul, Hermann Osthoff, Johannes Schmidt, Hugo Schuchardt, and Ferdinand de Saussure; their standpoints or methodologies are described in it clearly and succinctly. His more recent work, *In search for Indo-European Urheimat* (1993) is another exciting history pursuing the concept of an Indo-European homeland and its life and culture. He is also an author of two books concerning the Indo-European culture viewed from its vocabulary, namely, *Indo-European culture viewed from the aspect of words related to daily life* (1987) and *Indo-European culture from the aspect of words related to body parts and functioning* (1990). They belong in the tradition of academic works by Otto Schrader, Carl Darling Buck, and Emile Benveniste. In 1977 he had already translated into Japanese *Die Indogermanen* by Otto Schrader and Hans Krahe. His Latin-Greek 'contrastive' grammar, *A primary guide to Latin and Greek* (1998) is also a very useful handbook for younger scholars and students of linguistics.

Gengo Kenkyu, the journal of the Linguistic Society of Japan, has published his various linguistic

works, including 'Recent trends in Indo-European laryngeal theory' (1964: 45).

See also: Greek, Ancient; Kodzu, Harushige (1908–1973); Latin.

Bibliography

Delbrück B (1889). *Die indogermanischen Verwandtschaftsnamen.* Leipzig: Hirzel.

Kazama K (1964). 'Recent trends in Indo-European laryngeal theory.' *Gengo Kenkyu 45,* 27–38.

Kazama K (1977). *Die Indogermanen.* (Leipzig: Quelle & Meyer, 1935) Schrader O & Krahe H (translation into Japanese with notes).Tokyo: Kuronosu.

Kazama K (1978). *The birth of linguistics.* Tokyo: Iwanami Shoten.

Kazama K (1981). '*dwei-"zwei" im Altindischen.' *Working Papers in Linguistics 81,* 1–6.

Kazama K (1984). *Studies in Indo-European kinship terms.* Tokyo: Iwanami Shoten.

Kazama K (1986). 'A phase of the history of Indic.' *Journal of Indo-European Studies 14,* 289–300.

Kazama K (1987). *Indo-European culture viewed from the aspect of words related to daily life.* Tokyo: Heibonsha.

Kazama K (1990). *Indo-European culture from the aspect of words related to body parts and functioning.* Tokyo: Heibonsha.

Kazama K (1993). *In search for Indo-European Urheimat.* Tokyo: Iwanami Shoten.

Kazama K (1998). *A primary guide to Latin and Greek.* Tokyo: Sanseido.

Kempelen, Wolfgang von (1734–1804)

K-Å Forsgren, Umeå Universitet, Umeå, Sweden

Wolfgang von Kempelen was born in Bratislava (Pressburg) as the son of an official servant. He studied law and philosophy in Vienna, where he attracted the attention of Maria Theresa because of a translation of an act of law from Hungarian into German and was soon appointed Secretary and eventually Counselor of the Imperial Royal Court. Kempelen took a keen interest in physics and mechanics and also classical and modern languages and literature, writing poetry and authoring a drama. In other words, he was very creative and a man of versatile talent. Kempelen proved his skill as a constructor, as an organizer, and above all as an inventor of different useful devices. In his time he reached world-wide fame because of a less useful invention, a chess machine (the 'Turk'), which turned out to be a fraud: hidden inside it was a legless but skillful chess player! Kempelen died in Vienna at the age of 71.

The reason why Kempelen's name is still remembered in linguistics today, however, is his construction of a speaking machine, which was no joke but a very serious enterprise, on which he worked for 20 years without being able to finish the fourth version of it. The third version consisted, roughly, of a bellows (lungs), a compression box (bronchus), an ivory reed (vocal cords), a leather resonator (mouth), and two small pipes (nostrils). Bellow, resonator, and pipes were modified by means of the hands. In this way Kempelen was able to produce vowels and consonants, even short words and phrases.

The construction of the machine also resulted in a close study of the phonetic prerequisites of speech, which resulted in a book (*Mechanismus der menschlichen Sprache,* 1791) containing a thorough and, for his time, accurate description of speech organs and articulation, and also of his speaking machine. His introduction contains general reflections on human language and the origin of language. Basing his opinion on J. G. Herder, Kempelen argued against the idea that language is of divine origin as propagated, for example, by J. P. Süßmilch and A. Court de Guebelin. He was generally skeptical of speculation and showed a soundly critical mind. Referring to Hungarian and non-European languages, he not only rejected the idea of one proto-language of man but also the idea of a common origin of the other European languages. This is far from the standpoint taken by J. Grimm *et al.*, but Kempelen also saw the possibility of substrates ("*Vermischung der Völker*") and the necessity to consider not only words but also syntax and word order, which are less prone to change, by the comparison of languages.

In his passion for mechanical devices and automates, Kempelen was significant for his time and a typical exponent of the Cartesian (mechanistic) world conception. His speaking machine was unique, although there had been a simultaneous experiment in 1780 by C. G. Kratzenstein, St. Petersburg, whose device, however, could only produce vowels. Kempelen was also the first to understand that the mouth cavity and not, which had previously been assumed, the larynx was important for the quality of vowels. His work was developed further theoretically and practically up to the computerized speech synthesizer technique of today.

works, including 'Recent trends in Indo-European laryngeal theory' (1964: 45).

See also: Greek, Ancient; Kodzu, Harushige (1908–1973); Latin.

Bibliography

Delbrück B (1889). *Die indogermanischen Verwandtschaftsnamen.* Leipzig: Hirzel.

Kazama K (1964). 'Recent trends in Indo-European laryngeal theory.' *Gengo Kenkyu 45,* 27–38.

Kazama K (1977). *Die Indogermanen.* (Leipzig: Quelle & Meyer, 1935) Schrader O & Krahe H (translation into Japanese with notes).Tokyo: Kuronosu.

Kazama K (1978). *The birth of linguistics.* Tokyo: Iwanami Shoten.

Kazama K (1981). '*dwei-"zwei" im Altindischen.' *Working Papers in Linguistics 81,* 1–6.

Kazama K (1984). *Studies in Indo-European kinship terms.* Tokyo: Iwanami Shoten.

Kazama K (1986). 'A phase of the history of Indic.' *Journal of Indo-European Studies 14,* 289–300.

Kazama K (1987). *Indo-European culture viewed from the aspect of words related to daily life.* Tokyo: Heibonsha.

Kazama K (1990). *Indo-European culture from the aspect of words related to body parts and functioning.* Tokyo: Heibonsha.

Kazama K (1993). *In search for Indo-European Urheimat.* Tokyo: Iwanami Shoten.

Kazama K (1998). *A primary guide to Latin and Greek.* Tokyo: Sanseido.

Kempelen, Wolfgang von (1734–1804)

K-Å Forsgren, Umeå Universitet, Umeå, Sweden

Wolfgang von Kempelen was born in Bratislava (Pressburg) as the son of an official servant. He studied law and philosophy in Vienna, where he attracted the attention of Maria Theresa because of a translation of an act of law from Hungarian into German and was soon appointed Secretary and eventually Counselor of the Imperial Royal Court. Kempelen took a keen interest in physics and mechanics and also classical and modern languages and literature, writing poetry and authoring a drama. In other words, he was very creative and a man of versatile talent. Kempelen proved his skill as a constructor, as an organizer, and above all as an inventor of different useful devices. In his time he reached world-wide fame because of a less useful invention, a chess machine (the 'Turk'), which turned out to be a fraud: hidden inside it was a legless but skillful chess player! Kempelen died in Vienna at the age of 71.

The reason why Kempelen's name is still remembered in linguistics today, however, is his construction of a speaking machine, which was no joke but a very serious enterprise, on which he worked for 20 years without being able to finish the fourth version of it. The third version consisted, roughly, of a bellows (lungs), a compression box (bronchus), an ivory reed (vocal cords), a leather resonator (mouth), and two small pipes (nostrils). Bellow, resonator, and pipes were modified by means of the hands. In this way Kempelen was able to produce vowels and consonants, even short words and phrases.

The construction of the machine also resulted in a close study of the phonetic prerequisites of speech, which resulted in a book (*Mechanismus der menschlichen Sprache,* 1791) containing a thorough and, for his time, accurate description of speech organs and articulation, and also of his speaking machine. His introduction contains general reflections on human language and the origin of language. Basing his opinion on J. G. Herder, Kempelen argued against the idea that language is of divine origin as propagated, for example, by J. P. Süßmilch and A. Court de Guebelin. He was generally skeptical of speculation and showed a soundly critical mind. Referring to Hungarian and non-European languages, he not only rejected the idea of one proto-language of man but also the idea of a common origin of the other European languages. This is far from the standpoint taken by J. Grimm *et al.*, but Kempelen also saw the possibility of substrates ("*Vermischung der Völker*") and the necessity to consider not only words but also syntax and word order, which are less prone to change, by the comparison of languages.

In his passion for mechanical devices and automates, Kempelen was significant for his time and a typical exponent of the Cartesian (mechanistic) world conception. His speaking machine was unique, although there had been a simultaneous experiment in 1780 by C. G. Kratzenstein, St. Petersburg, whose device, however, could only produce vowels. Kempelen was also the first to understand that the mouth cavity and not, which had previously been assumed, the larynx was important for the quality of vowels. His work was developed further theoretically and practically up to the computerized speech synthesizer technique of today.

See also: Adelung, Johann Christoph (1732–1806); Grimm, Jacob Ludwig Carl (1785–1863); Herder, Johann Gottfried (1744–1803); Ihre, Johan (1707–1780).

Bibliography

Flanagan J L (1972). *Speech analysis, synthesis and perception.* Berlin: Springer Verlag.
Gessinger J (1994). *Auge und Ohr: Studien zur Erforschung der Sprache am Menschen 1700–1850.* Berlin: de Gruyter.
Kempelen W von (1970). *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine.* (*Faksimile-Neudruck der Ausgabe Wien 1791 mit einer Einleitung von Herbert E. Brekle und Wolfgang Wildgen.*) Stuttgart-Bad Cannstadt: Fromman.
Ungeheuer G (1962). *Elemente einer akustischen Theorie der Vokalartikulation.* Berlin: Springer Verlag.

Kennedy, Benjamin Hall (1804–1889)

V Van Elst, Seminarium Historiographiae Linguisticae, Leuven, Belgium

Benjamin Hall Kennedy was a distinguished classical scholar of wide reading, with a great talent for composition in Greek, Latin, and English verse, as can be seen, for example, in his work *Between whiles* (1877). He also edited several classical authors. His greatest contribution to the study of classics in England, however, was the publication of several successful school grammars, beginning with the *Latinae grammaticae curriculum, or a progressive grammar of the Latin language* in 1844.

Born at Summer Hill, near Birmingham, Kennedy spent his youth at Shrewsbury School, where he soon developed into a brilliant student. In 1823, he went to St John's College, Cambridge, where he gained several university awards. On his graduation Bachelor of Arts as 'Senior Classic' in 1827, he was elected fellow and classical lecturer of St John's College in 1828. Eight years later, he succeeded Dr Samuel Butler as headmaster of Shrewsbury (1836–1866). The last 22 years of his life he was Regius Professor of Greek at Cambridge.

In his lifetime, Kennedy was considered the greatest classical teacher of his century (Page, 1937–1938: 1303) for his powerful and enthusiastic manner of teaching, giving his pupils a lasting love for the classics. Today he is chiefly remembered for his grammar *The revised Latin primer* (1888), still widely used after revisions in 1909, 1930, and 1962. Its predecessor, *The public school Latin primer*, was published in 1866. It was commissioned by the headmasters of the nine chief English public schools – Shrewsbury included – with the encouragement of the Royal Commission of 1861, which recommended the use of a common Greek and Latin grammar in public schools. As the commissioners had selected the *Latinae grammaticae curriculum* as the basis of the new Latin grammar, Kennedy was asked to write the didactic manual. Subsequently, he also produced an advanced grammar as a supplement, *The public school Latin grammar* (1871), followed by an elementary version of the *Primer* entitled *The shorter Latin primer* (1888).

The *Primer* received a lot of criticism. Nevertheless, it rapidly became the standard Latin grammar in the public schools of England. Among its innovations was the introduction of the new case-order NOM (– VOC) – ACC – GEN – DAT – ABL, which, although adopted from – in addition to the Dane J. N. Madvig – the English grammarians T. H. Key and H. J. Roby (Kennedy, 1866: iv), remained linked in the English tradition with Kennedy's grammar and became the standard sequence in Britain (Allen-Brink, 1980).

See also: Latin; Madvig, Johann Nicolai (1804–1886).

Bibliography

Allen W S & Brink C O (1980). 'The old order and the new: a case history.' *Lingua 50,* 61–100.
Kennedy B H (1844). *Latinae grammaticae curriculum, or a progressive grammar of the Latin language.* London: Longmans, Green.
Kennedy B H (1866). *The public school Latin primer.* London: Longmans, Green.
Kennedy B H (1871). *The public school Latin grammar.* London: Longmans, Green.
Kennedy B H (1877). *Between whiles, or wayside amusements of a working life* (1st edn.). London: Bell.
Kennedy B H (1888). *The shorter Latin primer.* London: Longmans, Green.
Kennedy B H (1888). *The revised Latin primer.* London: Longmans, Green.

See also: Adelung, Johann Christoph (1732–1806); Grimm, Jacob Ludwig Carl (1785–1863); Herder, Johann Gottfried (1744–1803); Ihre, Johan (1707–1780).

Bibliography

Flanagan J L (1972). *Speech analysis, synthesis and perception.* Berlin: Springer Verlag.
Gessinger J (1994). *Auge und Ohr: Studien zur Erforschung der Sprache am Menschen 1700–1850.* Berlin: de Gruyter.
Kempelen W von (1970). *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine.* (*Faksimile-Neudruck der Ausgabe Wien 1791 mit einer Einleitung von Herbert E. Brekle und Wolfgang Wildgen.*) Stuttgart-Bad Cannstadt: Fromman.
Ungeheuer G (1962). *Elemente einer akustischen Theorie der Vokalartikulation.* Berlin: Springer Verlag.

Kennedy, Benjamin Hall (1804–1889)

V Van Elst, Seminarium Historiographiae Linguisticae, Leuven, Belgium

Benjamin Hall Kennedy was a distinguished classical scholar of wide reading, with a great talent for composition in Greek, Latin, and English verse, as can be seen, for example, in his work *Between whiles* (1877). He also edited several classical authors. His greatest contribution to the study of classics in England, however, was the publication of several successful school grammars, beginning with the *Latinae grammaticae curriculum, or a progressive grammar of the Latin language* in 1844.

Born at Summer Hill, near Birmingham, Kennedy spent his youth at Shrewsbury School, where he soon developed into a brilliant student. In 1823, he went to St John's College, Cambridge, where he gained several university awards. On his graduation Bachelor of Arts as 'Senior Classic' in 1827, he was elected fellow and classical lecturer of St John's College in 1828. Eight years later, he succeeded Dr Samuel Butler as headmaster of Shrewsbury (1836–1866). The last 22 years of his life he was Regius Professor of Greek at Cambridge.

In his lifetime, Kennedy was considered the greatest classical teacher of his century (Page, 1937–1938: 1303) for his powerful and enthusiastic manner of teaching, giving his pupils a lasting love for the classics. Today he is chiefly remembered for his grammar *The revised Latin primer* (1888), still widely used after revisions in 1909, 1930, and 1962. Its predecessor, *The public school Latin primer*, was published in 1866. It was commissioned by the headmasters of the nine chief English public schools – Shrewsbury included – with the encouragement of the Royal Commission of 1861, which recommended the use of a common Greek and Latin grammar in public schools. As the commissioners had selected the *Latinae grammaticae curriculum* as the basis of the new Latin grammar, Kennedy was asked to write the didactic manual. Subsequently, he also produced an advanced grammar as a supplement, *The public school Latin grammar* (1871), followed by an elementary version of the *Primer* entitled *The shorter Latin primer* (1888).

The *Primer* received a lot of criticism. Nevertheless, it rapidly became the standard Latin grammar in the public schools of England. Among its innovations was the introduction of the new case-order NOM (– VOC) – ACC – GEN – DAT – ABL, which, although adopted from – in addition to the Dane J. N. Madvig – the English grammarians T. H. Key and H. J. Roby (Kennedy, 1866: iv), remained linked in the English tradition with Kennedy's grammar and became the standard sequence in Britain (Allen-Brink, 1980).

See also: Latin; Madvig, Johann Nicolai (1804–1886).

Bibliography

Allen W S & Brink C O (1980). 'The old order and the new: a case history.' *Lingua 50*, 61–100.
Kennedy B H (1844). *Latinae grammaticae curriculum, or a progressive grammar of the Latin language.* London: Longmans, Green.
Kennedy B H (1866). *The public school Latin primer.* London: Longmans, Green.
Kennedy B H (1871). *The public school Latin grammar.* London: Longmans, Green.
Kennedy B H (1877). *Between whiles, or wayside amusements of a working life* (1st edn.). London: Bell.
Kennedy B H (1888). *The shorter Latin primer.* London: Longmans, Green.
Kennedy B H (1888). *The revised Latin primer.* London: Longmans, Green.

Page T E (1937–1938). 'Kennedy, Benjamin Hall, D. D.' In Stephen L & Lee S (eds.) *The dictionary of national biography X: Howard-Kenneth,* 3rd edn. London: Oxford University Press. 1302–1304.
Stray C (1994). 'The smell of Latin grammar: contrary imaginings in English classrooms.' *Bulletin of the John Rylands University Library of Manchester 76.3,* 201–220.
Stray C (1996). 'Primers, publishing, and politics: the classical textbooks of Benjamin Hall Kennedy.' *The Papers of the Bibliographical Society of America 90,* 451–474.
Stray C (ed.) (1999). *Classics in 19th and 20th century Cambridge. Curriculum, culture and community.* Cambridge: Cambridge University Press.

Kenya: Language Situation

C Githiora, University of London, London, UK

Typology

The Niger-Congo phylum is the largest. According to Heine and Mohlig (1980), approximately 65% of Kenyans speak one Bantu language or another (e.g., Swahili, Luhya (Luyia), Kamba, etc.), while about 3% of Kenyans speak a Cushitic language (e.g., Somali, Borana, Burji, etc.) belonging to the Afroasiatic phylum. The Nilo-Saharan family is represented by many Nilotic languages of the eastern Sudanic subgroup, which make up about 30% of the total languages spoken in Kenya (e.g., Masai [Maasi], Luo, Nandi, etc.). Each of the groups can further be broken into a number of dialect clusters. Area distribution of languages in Kenya makes a striking contrast. The majority Bantu languages are spoken in only about 20% of the national territory, while Nilotic languages occupy about 35% and the Cushitic languages more than 40% of Kenya's territory.

Kenya's Bantu languages are commonly divided into three general groups: western (Luyia, Kuria, etc.), central (Gikuyu, Kamba) and Coastal (Swahili, Digo, etc.). Guthrie's 1967–1971 classification places nearly all of them, except for Taveta (G21) and Swahili (G42), in Zone E. (SIL's [1990] Ethnologue lists the two as G20 and G40, respectively. It also lists Luyia and Bukusu in the J30 group, contrary to Guthrie's classification into E32 and E31c, respectively.) All languages except Swahili are tone languages, whereby distinctions in pitch are as important in a word as distinctions in consonants and vowels in indicating differences of meaning. Tone may also indicate grammatical differences. Some languages have two to four distinct tone levels while others have two, sometimes three levels plus a slight lowering of nonlow tones (downstep). In addition to tones, the chief characteristics of Bantu languages are noun classes whereby the head noun in a sentence or phrase 'agrees' with verbs and 'concords' with its modifiers such as adjectives, singular, and plural forms accordingly, as in the following highlighted examples from Swahili:

***M**ti **hu**u ni **m**refu*	'this tree is tall'
***mi**ti **hi**i ni **mi**refu*	'these trees are tall'
***m**toto **m**dogo **a**melala*	'the small child is sleeping/asleep'
***wa**toto **wa**dogo **wa**melala*	'the small children are sleeping/asleep'

There are no Khoisan languages found in Kenya. The chief, well-known characteristic of these languages are clicks, or consonant sounds that are found in the initial position of almost every noun, verb, and adjective. Immigrants who have settled in the country for a few generations maintain languages of the Indo-European family alive, such as Punjabi (Punjabi, Eastern), Gujarati, Hindi, and English. There exists also a Kenyan Sign Language (KSL) for the deaf, approved by relevant official institutions such as the Kenya National Association of the Deaf. KSL appears mainly unrelated to other countries' sign languages, although it has been standardized with slight variations since 1961. It would appear, for example, that the deaf in neighboring Uganda and Tanzania do not really understand KSL, but those living in different parts of Kenya can understand each other completely even with some dialect differences. There is a manual alphabet for spelling and a dictionary.

Languages and Dialects

The common difficulties of establishing clear boundaries between languages and dialects are found in Kenya, which is why the estimate of total languages ranges from 30 to 60, according to different published reports. The problem is most acute when dealing with the numerous, contiguous Bantu languages spoken by two-thirds of the Kenyan population. More crucially, dialect boundaries are not clearly cut isoglosses, but rather dialect continua. In the case of Kenya, further obfustication is the result of a lack of distinction between language, ethnicity,

Page T E (1937–1938). 'Kennedy, Benjamin Hall, D. D.' In Stephen L & Lee S (eds.) *The dictionary of national biography X: Howard-Kenneth,* 3rd edn. London: Oxford University Press. 1302–1304.
Stray C (1994). 'The smell of Latin grammar: contrary imaginings in English classrooms.' *Bulletin of the John Rylands University Library of Manchester 76.3,* 201–220.
Stray C (1996). 'Primers, publishing, and politics: the classical textbooks of Benjamin Hall Kennedy.' *The Papers of the Bibliographical Society of America 90,* 451–474.
Stray C (ed.) (1999). *Classics in 19th and 20th century Cambridge. Curriculum, culture and community.* Cambridge: Cambridge University Press.

Kenya: Language Situation

C Githiora, University of London, London, UK

Typology

The Niger-Congo phylum is the largest. According to Heine and Mohlig (1980), approximately 65% of Kenyans speak one Bantu language or another (e.g., Swahili, Luhya (Luyia), Kamba, etc.), while about 3% of Kenyans speak a Cushitic language (e.g., Somali, Borana, Burji, etc.) belonging to the Afroasiatic phylum. The Nilo-Saharan family is represented by many Nilotic languages of the eastern Sudanic subgroup, which make up about 30% of the total languages spoken in Kenya (e.g., Masai [Maasi], Luo, Nandi, etc.). Each of the groups can further be broken into a number of dialect clusters. Area distribution of languages in Kenya makes a striking contrast. The majority Bantu languages are spoken in only about 20% of the national territory, while Nilotic languages occupy about 35% and the Cushitic languages more than 40% of Kenya's territory.

Kenya's Bantu languages are commonly divided into three general groups: western (Luyia, Kuria, etc.), central (Gikuyu, Kamba) and Coastal (Swahili, Digo, etc.). Guthrie's 1967–1971 classification places nearly all of them, except for Taveta (G21) and Swahili (G42), in Zone E. (SIL's [1990] Ethnologue lists the two as G20 and G40, respectively. It also lists Luyia and Bukusu in the J30 group, contrary to Guthrie's classification into E32 and E31c, respectively.) All languages except Swahili are tone languages, whereby distinctions in pitch are as important in a word as distinctions in consonants and vowels in indicating differences of meaning. Tone may also indicate grammatical differences. Some languages have two to four distinct tone levels while others have two, sometimes three levels plus a slight lowering of nonlow tones (downstep). In addition to tones, the chief characteristics of Bantu languages are noun classes whereby the head noun in a sentence or phrase 'agrees' with verbs and 'concords' with its modifiers such as adjectives, singular, and plural forms accordingly, as in the following highlighted examples from Swahili:

***M**ti h**uu** ni **m**refu*	'this tree is tall'
***m**iti h**ii** ni **m**irefu*	'these trees are tall'
***m**toto **m**dogo **a**melala*	'the small child is sleeping/ asleep'
***wa**toto **wa**dogo **wa**melala*	'the small children are sleeping/ asleep'

There are no Khoisan languages found in Kenya. The chief, well-known characteristic of these languages are clicks, or consonant sounds that are found in the initial position of almost every noun, verb, and adjective. Immigrants who have settled in the country for a few generations maintain languages of the Indo-European family alive, such as Punjabi (Punjabi, Eastern), Gujarati, Hindi, and English. There exists also a Kenyan Sign Language (KSL) for the deaf, approved by relevant official institutions such as the Kenya National Association of the Deaf. KSL appears mainly unrelated to other countries' sign languages, although it has been standardized with slight variations since 1961. It would appear, for example, that the deaf in neighboring Uganda and Tanzania do not really understand KSL, but those living in different parts of Kenya can understand each other completely even with some dialect differences. There is a manual alphabet for spelling and a dictionary.

Languages and Dialects

The common difficulties of establishing clear boundaries between languages and dialects are found in Kenya, which is why the estimate of total languages ranges from 30 to 60, according to different published reports. The problem is most acute when dealing with the numerous, contiguous Bantu languages spoken by two-thirds of the Kenyan population. More crucially, dialect boundaries are not clearly cut isoglosses, but rather dialect continua. In the case of Kenya, further obfustication is the result of a lack of distinction between language, ethnicity,

and culture. Nonlinguistic factors (e.g., prestige, historic-cultural ties, etc.) in a country where ethnolinguistic nationalism is prevalent may at times overshadow strict linguistic criteria normally applied to demarcate language and dialect boundaries. Where for example, Masaba, Luyia, and Bukusu are treated as different languages, linguistically this may not be so readily evident, as mutual intelligibility among them is quite high. Conversely, less mutually intelligible languages such as a Kipsikiis and Nandi are considered dialects of the same language, Kalenjin. Population data and censuses of the past have used ethnic rather than linguistic affiliation to identify people. For example, the northeastern part of the country is considered Somali-speaking because it is seen as little differentiated culturally. Yet, in addition to the more numerous Somali speakers, there are a number of distinct Oromo (Cushitic) languages such as Borana, Boni, and Garreh-Ajuran. People often identify as Luhya or Kalenjin, terms that cover several differentiated dialects and even languages. Many may also identify as Swahili because of religious and cultural practices, without letting it be known that their vernacular language is really different.

Swahili

Swahili language occupies a special position in Kenya's linguistic landscape. It is the national and official language of Kenya, now a part of the new draft constitution (2003) as such. Linguistically, it is of the coastal Bantu subgroup (G40), with several distinctive dialects spoken on the Indian Ocean islands of Lamu, Pate, Pemba, and Zanzibar. The language developed to become the common coastal language as early as the 13th century and has been greatly influenced by Arabic, with some loan words from Hindi, Persian, Portuguese, and English. Its capability to adapt to changing contexts has contributed to its unique position among the languages of East and Central Africa. Swahili is the universal lingua franca of Kenya today in small-scale trade and media. It is closely connected with urban life and with certain occupations. Our recent survey of language use in Nairobi revealed, for example, that although only 7% of the total number of respondents ($n = 1$, 300) reported Swahili as a first language, more than 50% use Swahili in their household interactions with other family members (Githiora, 2002). Swahili is in theory used as a medium of instruction in urban schools, but in practice English plays a larger role, according to teachers who participated in the same survey. However, Swahili is taught as a subject at all levels of education in Kenya, with many graduates specializing in Swahili each year from the three main national universities. Radio and television broadcasts may be heard in Swahili, English, and various Kenyan languages. In rural areas of Kenya, Swahili is usually encountered only in radio and television broadcasts, shops or business run by nonlocals, or in Swahili newspapers and pamphlets. In rural schools, children are taught in their local language up to Standard 4, but are taught Swahili as well as English as a second, examinable language. Overall, about two-thirds of Kenyans can make use of Swahili as opposed to less than one-sixth who know English (Heine and Mohlig, 1980: 61). The relationship between language, power, and ideology becomes highly relevant, as well the history and trajectory of Swahili during the modern, colonial, and post-colonial periods. For example, the language is closely connected with ethnic neutrality, urban life, and with certain occupations. Consequently, its functional roles within Kenyan society are strongly linked to the rural–urban divide, and to socioeconomic class. In addition to its geographical dialects, at least one distinct urban Swahili-based mixed code known as Sheng (Githiora, 2002) has emerged since the time of independence in the 1960s, and its influence is being felt far in the rural areas as well. Growing attention to this urban, mixed code is important in so far as how the code affects communication, education, and the desired spread and use of standard Swahili in Kenya.

Multilingualism and Language Policy

There is a high level of multilingualism among Kenyans, as citizens speak at least one other second language in addition to their first. Put differently by Heine and Mohlig (1980), Kenyans have on average a second language competence of 1.01 languages. A recent study of multilingualism and language use in Nairobi city, a microcosm of Kenya's linguistic mosaic, showed that the average number of languages spoken by a city resident is actually three, while a 25% of the total sampled live in bilingual households (Githiora, 2002). The higher mean figure in the findings in Nairobi city are probably because urban dwellers tend to speak both Swahili and English, unlike rural Kenyans. Kenya's language policy is best defined as nonexistent, despite a 1969 declaration by the ruling party (KANU) in parliament at that time, whereby Swahili was to become the official language of the country within a 5-year period. This declaration was made in good spirit, but it lacked the backing of any serious research on Kenya's sociolinguistic situation, nor was there a defined method in implementing the proposal. There was no scientific basis for defining the different roles that each or some of the 50 odd languages and their dialects could play

in different tasks within a modern Kenyan state. Such a sociolinguistic understanding would not come about until the 1980 publication of the *Atlas of Kenya's languages and dialects* by Heine and Mohlig, but even so, no concrete policy proposals have emerged backed by this ambitious study. Today, bills, parliamentary debates, laws, and statutes are still drafted in English, but parliamentarians who wish to can debate in Swahili while court interpreters a provide translation service into Swahili or any other Kenyan language as required, including Kenyan Sign Language (KSL) for the deaf.

Minority and Endangered Languages

While Swahili dominates study and discussion of Kenya's languages, a number of minority languages occupy an important position as they relate to issues of endangered languages, language contact, shift, and death. Among the larger groups of speakers, a form of stable bilingualism exists where, for example, two-thirds of Kenyans can speak Swahili, and a far smaller number are competent in English without losing competence in their vernacular languages. The situation is different in the case of smaller ethnolinguistic communities whereby bilingualism tends to be of the assimilating type. One of the consequences of cultural assimilation and adoption of the lifestyle of larger, economically better off communities is language shift which is seen as a necessity to adapt to the new value system. In most cases, it implies major changes of lifestyle such as a shift from fishing to pastoralism as in case of El Molo, Kenya's most endangered language whose estimated number of speakers left ranges from 8 to 500 on the shores of Lake Turkana in northern Kenya. This process culminating in a complete language shift has taken place (i.e., the languages are nearly extinct) among the Yaaku and Omotik, two Cushitic language speaking communities in favor of Masai, a Nilotic language. A similar process of cultural and linguistic assimilation by Gikuyu-, Kalenjin-, and Masai-speaking communities also places Okiek, a Nilotic language second on the list of Kenyan near-extinct languages. It is telling that the three assimilating communities mentioned above refer to the Okiek people and language as Ndorobo. Another case is that of Suba, a Bantu language that is also nearly completely absorbed by the larger, numerically dominant Luo (Nilotic) language. No significant pidgins have emerged from Kenya. What has been termed as Cutchie Swahili (Asian Swahili) or Shamba Swahili spoken by Asian immigrant traders and European settler farmers respectively is basically a nonstandard Swahili sociolect. The only true pidgin language of Kenya is Nubi (or Kinubi), a creolized form of Arabic spoken by a small minority of people whose origins date back to the colonial period in northern Uganda and Sudan, who now live in two communities; the larger one in Nairobi city's Kibera quarter, and a smaller one a few kilometers from Kisii town in western Kenya.

See also: Applied Linguistics in Africa; Bilingualism and Second Language Learning; Contact-Induced Convergence: Typology and Areality; Education in a Multilingual Society; Endangered Languages; Gikuyu; Identity: Second Language; Language Change and Language Contact; Language Education: Grammar; Language Education Policies in Africa; Language Policy in Multilingual Educational Contexts; Language Surveys; Lingua Francas as Second Languages; Luo; Sign Languages of the World; Swahili; Tanzania: Language Situation; Tone: Phonology.

Language Maps (Appendix 1): Map 17.

Bibliography

Ethnologue (2004). *Ethnologue: Languages of the world* (14th edn.). Dallas, TX: SIL International.

Githiora C (2002). 'Sheng: peer language, Swahili dialect or emerging creole?' *Journal of African Cultural studies 15(2),* 159–181.

Heine B & Möhlig W (1980). *The atlas of the languages and dialects of Kenya.* Berlin: Dietrich Reimer Verlag.

Whiteley W (ed.) (1974). *Language in Kenya.* Nairobi, Kenya: Oxford University Press.

Relevant Website

http://www.ethnologue.com – Ethnologue.

Keres

M J Mixco, University of Utah, Salt Lake City, UT, USA

Introduction

Of the 21 Pueblo tribes in the American Southwest, one speaks Zuni, a language isolate; another speaks Hopi (Uto-Aztecan family) in several Arizona villages. In New Mexico, Tanoan pueblos (speaking Tiwa, Tewa, or Towa, related to Kiowa) are interspersed with those speaking Keres (or Keresan, *Quirix* or *Quires* in 16th-century Spanish; Davis, 1959).

Historically, Keres borrowed lexically from other pueblo languages via bilingual speakers, now a defunct process. Keresan has two major dialect divisions: Eastern (Rio Grande Valley) and Western (Colorado River drainage). Mutual intelligibility among dialects increases with proximity (speaker populations follow in parentheses as cited in Mithun, 1999, from Valiquette, 1995), Western: Acoma (1930)/Laguna (2060); Eastern: Zia (504)/Santa Ana (384)-San Felipe (1985)/Santo Domingo (2965)-Cochití (525).

Beginning in 1598, Spanish influences were strong only in the area of material culture (livestock, new crops, trade goods, and technologies). The impact was slighter in nonmaterial domains (e.g., religion) because of indigenous resistance resulting in the Pueblo Revolt of 1680–1692.

Surrounded by denser Spanish populations, the Eastern dialects received the most loanwords via Spanish bilingualism that persisted into the mid-20th century. English displaced Spanish as the regional lingua franca but provided fewer loans. Keres linguistic literature is scant, despite intensive Pueblo ethnographic work.

Structural Overview

Keres consonantal series are plain (b, d, dʸ, g), aspirated (p, t, ʧ, k), and glottalized (p', t', ʧ', k') stops and affricates; sonorants (m, n, r, w, y, m', n', r', w', y'); and fricatives (s, ʂ, ʃ, s', ʂ', ʃ'). There are two laryngeals (ʔ, h); five long and short vowels, voiced and voiceless contextually (i, e, a, ɨ, u); and as many as four pitch accents (high ´, falling ˆ, breathy ˋ, and glottal ‘), depending on the dialect. Morphologically, Keres is polysynthetic, exhibiting distinct sets of prefixes on active and stative verbs to express nominal arguments without independent pronouns. Pronominal prefixes, with the subject and object fused into one morpheme, undergo especially complex phonological alternations and simultaneously encode five subject and object persons, including an indefinite and a fourth or obviative person (when the subject is hierarchically lower than the object). Also simultaneous is modality (negative, dubitative, hortative, negative hortative, future hortative, and indicative). Singular, dual, and plural numbers occur. Voice requires its own set of pronominal prefixes followed by the reflexive-reciprocal prefix /-a-/ or the passive /-aʔa- or –a'-/. Benefactives are expressed discontinuously by prefixes and suffixes. Aspect (continuous, fulfilled, and state) is suffixal, as is adverbial subordination. Some intransitive pronominal agreement examples are [ʂ̥-u'pɨ̥] ‘I or you ate,’ [k-u'pa] ‘he ate’ versus the transitives [s'-ˋaku] ‘I bit you’, [ʧ'-ˋaku] ‘you bit me,’ [ʂ̥-éku̥] ‘I bit (him),’ [s'-ˋaku] ‘he bit (him),’ [sg-àku] ‘someone bit him’ (Mithun, 1999: 438–440). The following sentences are from Davis (1964) as quoted in Mithun (1999):

ʔe ʂu ʔe yúsi n'ɨ ʔeu dya:mi ʔeu ʂu ʔe yúsi n'ɨ t'-a'gúyan-e
from there down EMPH eagle from there down 3.DUB-PASS-send-PL
‘then the eagle was sent down from above’

ʂu ʔe hauʔ dí-uw'ác'i
PRT PRT toward 3.DUB-approach
‘(the eagle) approached’

Language, Culture, and Society

The Keres language reflects the traditional interrelated cultural concerns of the Pueblo peoples, principally religious ceremonialism, agriculture, and theocratic governance. This has given rise to elaborate esoteric terminologies distinguishing the realms of the sacred from the profane. The latter categories may vary from pueblo to pueblo. There is baby talk as well as a difference between male and female speech in frequently occurring words (Kroskrity, 1983; Sims and Valiquette, 1990).

Recently, there have been vigorous efforts in some Keresan pueblos to counteract language shift (Pecos and Blum-Martinez, 2001; Sims, 2001).

See also: Hopi; United States of America: Language Situation.

Bibliography

Campbell L (1997). *American Indian languages: the historical linguistics of Native America.* Chicago: University of Chicago Press.

Davis I (1959). 'Linguistic clues to northern Rio Grande prehistory.' *El Palacio 66,* 73–84.
Davis I (1963). 'Bibliography of Keresan linguistic sources.' *IJAL 29,* 289–293.
Davis I E (1964). 'The language of Santa Ana.' *Bureau of American Ethnology Bulletin 191,* 53–190.
Davis I E (1979). 'The Kiowa-Tanoan, Keresan, and Zuni languages.' In Campbell L & Mithun M (eds.) *The languages of Native America: historical and comparative assessment.* Chicago: University of Chicago Press.
Kroskrity P V (1983). 'On male and female speech in the Pueblo Southwest.' *IJAL 49,* 88–91.
Miller W R (1965). *Acoma grammar and texts* (vol. 40). University of California Publications in Linguistics, Berkeley: University of California Press.
Miller W R & Davis I E (1963). 'Proto-Keresan phonology.' *IJAL 29,* 310–330.
Mithun M (1999). *The languages of native North America.* Cambridge Language Surveys. Cambridge, UK: Cambridge University Press.
Pecos R & Blum-Martinez R (2001). 'The key to cultural survival: language planning and revitalization in the Pueblo de Cochití.' In Hinton L & Hale K (eds.) *The green book of language revitalization in practice.* New York: Academic Press.
Sims C P (2001). 'Native language planning: a pilot process in the Acoma Pueblo community.' In Hinton L & Hale K (eds.) *The green book of language revitalization in practice.* New York: Academic Press.
Sims C P & Valiquette H (1990). 'More on male and female speech in (Acoma and Laguna) Keresan.' *IJAL 56,* 162–166.
White L A (1944). 'A ceremonial vocabulary among the Pueblos.' *IJAL 10,* 161–167.
Valiquette H P (1991). 'A study for a lexicon of Laguna Keresan.' *Dissertation Abstracts International, A: The Humanities and Social Sciences 52(2),* 527-A.

Ket

G D S Anderson, University of Oregon, Eugene, OR, USA

Ket [KET] is also known as Yenisei Ostyak and Imbat[skij] Ket. There are probably fewer than 200 mother tongue speakers out of an ethnic population of 1100–1200 (Krivonogov, 1998). Ket is spoken in north central Siberia along the Yenisei river and its tributaries (e.g., Yelogui), in northern Krasnoyarski Kray, spoken mainly in the villages of Sulomai, Kellog, Surgutikha, and Maiduka. Ket is the sole surviving member of the Yeniseic language family. Related languages include Arin, Assan, and Pumpokol, which became extinct by mid-18th century; Kott, which lasted until the mid-19th century, originally spoken to the north, west, and south of Krasnoyarsk. The closely related Yugh (also known as Sym-Ket) became extinct only in 1980s, and was last spoken in Vorogovo village. Yeniseic language speakers shifted to Russian or various Turkic varieties, in particular Chulym Turkic, Xakas, Shor, and Northern Altai varieties. Yeniseic speakers previously occupied the territory south along the Yenisei to the mouth of the Dupches River. Yeniseic-speaking peoples once occupied a large area in western and central Siberia, based on the widespread use of Yeniseic hydronymics across the area. Presumably, the attested Yeniseic peoples were encroached upon and marginalized areally by the advance at various periods of Samoyedic- and Ob-Ugric-speaking peoples from the west, and Turkic and Tungusic from the south and east, until they occupied their attested position along the Yenisei. The Xiong Nu of the Chinese chronicles may have spoken a Yeniseic language.

There are three dialects of Ket: Northern, Central, Southern. Mainly Southern Ket survives in such villages as Kellog and Sulomaj. Ket is severely endangered but enjoys high status in the villages where Ket people dominate. Only a handful of young people are learning the language as a first language, but Ket instruction in primary schools has begun in certain villages (e.g., Kellog), based on the written form of Ket developed by native Ket scholars and Professor Heinrich Werner.

Ket is unique among central Siberian languages for its unusual system of tone. Tones appear as a prosodic feature of the two leftmost syllables (if present) in Ket words (Vajda, 1999).

(1) Southern Ket
$^{1}s^{j}ul^{j}$ 'blood'
$^{2}s^{j}u^{ʔ}l^{j}$ 'white salmon'
$^{3}s^{j}u{:}l^{j}$ 'sled'
$^{4}s^{j}ul^{j}$ 'cradle hook'
(Vajda, 1999: 5)

Tone differentiates both lexical and grammatical forms in Ket.

(2) Southern Ket
ásɛ̀l 'ski' *àsɛ́l* 'type of large covered houseboat'

Davis I (1959). 'Linguistic clues to northern Rio Grande prehistory.' *El Palacio 66,* 73–84.
Davis I (1963). 'Bibliography of Keresan linguistic sources.' *IJAL 29,* 289–293.
Davis I E (1964). 'The language of Santa Ana.' *Bureau of American Ethnology Bulletin 191,* 53–190.
Davis I E (1979). 'The Kiowa-Tanoan, Keresan, and Zuni languages.' In Campbell L & Mithun M (eds.) *The languages of Native America: historical and comparative assessment.* Chicago: University of Chicago Press.
Kroskrity P V (1983). 'On male and female speech in the Pueblo Southwest.' *IJAL 49,* 88–91.
Miller W R (1965). *Acoma grammar and texts* (vol. 40). University of California Publications in Linguistics, Berkeley: University of California Press.
Miller W R & Davis I E (1963). 'Proto-Keresan phonology.' *IJAL 29,* 310–330.
Mithun M (1999). *The languages of native North America.* Cambridge Language Surveys. Cambridge, UK: Cambridge University Press.
Pecos R & Blum-Martinez R (2001). 'The key to cultural survival: language planning and revitalization in the Pueblo de Cochití.' In Hinton L & Hale K (eds.) *The green book of language revitalization in practice.* New York: Academic Press.
Sims C P (2001). 'Native language planning: a pilot process in the Acoma Pueblo community.' In Hinton L & Hale K (eds.) *The green book of language revitalization in practice.* New York: Academic Press.
Sims C P & Valiquette H (1990). 'More on male and female speech in (Acoma and Laguna) Keresan.' *IJAL 56,* 162–166.
White L A (1944). 'A ceremonial vocabulary among the Pueblos.' *IJAL 10,* 161–167.
Valiquette H P (1991). 'A study for a lexicon of Laguna Keresan.' *Dissertation Abstracts International, A: The Humanities and Social Sciences 52(2),* 527-A.

Ket

G D S Anderson, University of Oregon, Eugene, OR, USA

Ket [KET] is also known as Yenisei Ostyak and Imbat[skij] Ket. There are probably fewer than 200 mother tongue speakers out of an ethnic population of 1100–1200 (Krivonogov, 1998). Ket is spoken in north central Siberia along the Yenisei river and its tributaries (e.g., Yelogui), in northern Krasnoyarski Kray, spoken mainly in the villages of Sulomai, Kellog, Surgutikha, and Maiduka. Ket is the sole surviving member of the Yeniseic language family. Related languages include Arin, Assan, and Pumpokol, which became extinct by mid-18th century; Kott, which lasted until the mid-19th century, originally spoken to the north, west, and south of Krasnoyarsk. The closely related Yugh (also known as Sym-Ket) became extinct only in 1980s, and was last spoken in Vorogovo village. Yeniseic language speakers shifted to Russian or various Turkic varieties, in particular Chulym Turkic, Xakas, Shor, and Northern Altai varieties. Yeniseic speakers previously occupied the territory south along the Yenisei to the mouth of the Dupches River. Yeniseic-speaking peoples once occupied a large area in western and central Siberia, based on the widespread use of Yeniseic hydronymics across the area. Presumably, the attested Yeniseic peoples were encroached upon and marginalized areally by the advance at various periods of Samoyedic- and Ob-Ugric-speaking peoples from the west, and Turkic and Tungusic from the south and east, until they occupied their attested position along the Yenisei. The Xiong Nu of the Chinese chronicles may have spoken a Yeniseic language.

There are three dialects of Ket: Northern, Central, Southern. Mainly Southern Ket survives in such villages as Kellog and Sulomaj. Ket is severely endangered but enjoys high status in the villages where Ket people dominate. Only a handful of young people are learning the language as a first language, but Ket instruction in primary schools has begun in certain villages (e.g., Kellog), based on the written form of Ket developed by native Ket scholars and Professor Heinrich Werner.

Ket is unique among central Siberian languages for its unusual system of tone. Tones appear as a prosodic feature of the two leftmost syllables (if present) in Ket words (Vajda, 1999).

(1) Southern Ket
$^{1}s^{j}ul^{j}$ 'blood'
$^{2}s^{j}u^{ʔ}l^{j}$ 'white salmon'
$^{3}s^{j}u{:}l^{j}$ 'sled'
$^{4}s^{j}ul^{j}$ 'cradle hook'
(Vajda, 1999: 5)

Tone differentiates both lexical and grammatical forms in Ket.

(2) Southern Ket
ásὲl 'ski' *àsέl* 'type of large covered houseboat'

bʌ̀ntán	'mallard duck'	*bʌ́ntàn*	'mallard ducks'

(Vajda, 1999: 13)

The Ket dialects differ tonally primarily in the realization of the fourth tone, which may be falling with a short vowel and no pharyngealization (Southern Ket), or pharyngealized with a long vowel (Central and Northern Ket).

(3)

Southern Ket	Central Ket	Northern Ket	gloss
^{4}*sʲɛlʲ*	^{4}*sʲɛ:lʲi*	^{4}*sʲɛ:lʲi*	'reindeer'
^{4}*asʲ*	^{4}*a:sʲe*	^{4}*a:sʲe*	'feather'
^{4}*ɨrʲ*	^{4}*ɨ:də*	^{4}*ɨ:rʲe*	'spring'

(Vajda, 1999: 7)

Although Ket is unique among the languages of Siberia for its tone/register system (certain Ket tones, e.g., tone 2, are associated with pharyngeal tension), Ket is typical of north-central Siberian languages for its elaborate case system (Anderson, 2003). Nouns may appear in one of ten or eleven case forms. Ket is unusual for Siberian languages in distinguishing three genders or noun classes, roughly masculine, feminine, and other/neuter. One set of the case forms encodes these distinctions. Others do not and appear to be more recently fused postpositional constructions.

(4) Ket

hɨb-daŋa	*hun-diŋa*
son-DATIVE.MASCULINE	daughter-DATIVE.FEMININE
'to (his) son'	'to (his) daughter'

(Werner, 1997: 105)

In addition, Ket makes extensive use of case forms on verbs as well within a highly diversified system of case-marked clausal subordination that dominates complex sentence structure in the languages of the region.

(5) *-diŋalʲ* Ablative → 'after,' 'since'

Ket

bu	*ətn-asʲ*	*du-ɣ-a-raq-diŋalʲ*
he	we-INS/COM	I-SEP-PRES-live-ABL
dɔŋ	*sɨkŋ*	*u-ɣɔñ*
three	year.PL	III-go

'three years have passed since he's been living with us'

(Werner, 1997: 353)

Among the most noteworthy features of Ket from a typological and areal perspective is its elaborate and highly complex verbal system. There are many different structural positions or slots within the Ket verb, the exact number of which is debated by Ket specialists (e.g., Werner (1997) assigns 18 such slots (14 prefix positions, a root, and three suffix positions), while Vajda (2003), on the other hand, considers there to be 10 such slots in the Ket verb (eight prefix position, root + one suffix position). Some examples of complex Ket verb forms include the following:

(6)

da-bágdɛ̀ŋ-ú-ɣ-à-vɛt	*d-us'-n-ba-ɣɔ-v-in'-tet*
3FEM-drag-3NEUT.OBJ-DIR-PRES-ITER	3M-INCORP-DER-1-PST-INAN-PST.PRF-hit
'she drags it often'	'he has hit me'
(Vajda, 2003: 63)	(Werner, 1997: 156)

Morphosyntactically, Ket stands out for its unusual predilection for multiple encoding of a single category, e.g., subject in verbs or plurality in nouns. Examples of this kind of redundant encoding may be seen in the following examples.

(7) Ket

d-dəŋ-b-it-n	^{1}*qa-ŋ-s-eŋ-naŋal'*
1-1PL-INAN-transport-PL	chief..-PL-..chief-PL-PL:ABL
'we transport it'	'from the chiefs'
(Shabaev, 1987; Werner, 1995)	
	< ^{4}qa[:]s 'chief'

A number of proposals have been offered on the possible wider genetic affiliation of the Yeniseic languages, including connections with Burushaski, Sino-Tibetan, and Northeast Caucasian languages of Eurasia, as well as Athabaskan (Na-Dene) languages of North America. To date, only the latter proposal has met with any positive reactions among specialists.

See also: Endangered Languages; Russian Federation: Language Situation.

Bibliography

Anderson G D S (1996). 'Ket tonal alternations in absolutive plural.' In Aronson H I (ed.) *Non-Slavic languages 8: linguistic studies, new series.* Chicago: Chicago Linguistic Society. 1–18.

Anderson G D S (2003). 'Yeniseic languages from a Siberian areal perspective.' In Vajda E & Anderson G D S (eds.) *Studia Yeniseica: typological studies on Yeniseic in honor of Heinrich Werner, Sprachtypologie und Universalienforschung 56(1/2).* München/Bremen: Akademie Verlag. 12–39.

Belimov E I (1991). *Ketskij sintaksis.* Novosibirsk.

Bibikova V S (1971). 'Grammaticheskie priznaki imeni prilagatel'nogo v ketskom jazyke.' *Jazyki i toponimija Sibiri 4,* 27–36.

Bouda K (1957). 'Die Sprache der Jenissejer: Genealogische und morphologische Untersuchungen.' *Anthropos 52,* 65–134.

Castrén M A (1858). *Versuch einer jenissej-ostjakischen und kottischen Sprachlehre nebst Wörterverzeichnissen aus den genannten Sprachen.* St. Petersburg.

Comrie B (1982). 'Verb agreement in Ket.' In Aronson H I (ed.) *Folia Slavica 5.1–3. Papers from the Second Conference on the non-Slavic Languages of the USSR*. Columbus, Ohio: Slavica. 115–127.

Denning R F (1976). 'Palatalizatsija soglasnykh v ketskom jazyke.' *Jazyki i toponimija Sibiri 6*, 85–89.

Donner K (1930). 'Über die Jenissei-Ostjaken und ihre Sprache.' *Journal de la Société Finno-Ougrienne 44*, 1–32.

Dul'zon A P (1961). 'Slovarnye materialy XVIII v. po enisejskim narechijam.' *Uchenye zapiski Tomskogo pedagogicheskogo instituta 19(2)*, 152–189.

Dul'zon A P (1966). *Ketskie skazki*. Tomsk.

Dul'zon A P (1968). *Ketskij jazyk*. Tomsk: Izdatel'stvo tomskogo universiteta.

Gajer R S (1981). 'Skazki i bytovye teksty.' In Porotova T I (ed.) *Skazki narodov sibirksogo Severa 4*. Tomsk. 27–40.

Grishina N M (1981). 'Slozhnopodchinennye predlozhenija.' In Porotova T I (ed.) *Skazki narodov sibirksogo Severa 4*. Tomsk. 92–121.

Karger N K (1937). 'Ketskij jazyk.' In *Jazyki i pis'mennost' narodov Severa III*. Leningrad: Krejnovich, È. A. 223–238.

Khelimskij E A (1986). 'Arkhivnye materialy XVIII v. po enisejskim jazykam.' In Ja Skorik P (ed.) *Paleoaziatskie jazyki*. Leningrad: Nauka. 179–212.

Kostjakov M M (1979). 'Vremja raskhozhdenija ketskogo i kottskogo jazykov po dannym leksikostatistiki.' In Ubrjatova E I (ed.) *Voprosy stroja enisejskikh jazykov*. Novosibirsk: SO Akademija Nauk SSSR. 118–127.

Kostjakov M M (1986). 'Slozhnoe slovo i slovosochetanie v ketskom jazyke.' In Vall M N (ed.) *Issledovanija po grammatike i leksike enisejskikh jazykov*. Novosibirsk: Nauka. 59–69.

Krejnovich E A (1968a). *Glagol ketskogo jazyka*. Leningrad: Nauka.

Krejnovich E A (1968b). 'Sposoby dejstvija v glagole ketskogo jazyka.' In Ivanov V V *et al.* (eds.) *Ketskij sbornik. Lingvistika*. Moscow: Moskovskij Gosudarstvennyj Universitet. 75–138.

Krejnovich E A (1969). 'Medvezhij prazdnik u ketov.' In Ivanov V V *et al.* (eds.) *Ketskij sbornik. Mifologija, ètnografija, teksty*. Moscow: Moskovskij Gosudarstvennyj Universitet. 6–112.

Krivonogov V P (1998). *Kety na poroge III tysjacheletija*. Krasnoyarsk: Krasnoyarskij pedagogicheskij universitet.

Maksunova Z (2003). 'Compounding and incorporation in Ket (central dialect).' In Vajda E J & Anderson G D S (eds.) *Studia Yeniseica: typological studies on Yeniseic in honor of Heinrich Werner, Sprachtypologie und Universalienforschung 56(1/2)*, 123–132.

Minaeva V P (2003). 'Russian grammatical interference in Ket.' In Vajda E J & Anderson G D S (eds.) *Studia Yeniseica: Typological Studies on Yeniseic in honor of Heinrich Werner, Sprachtypologie und Universalienforschung 56(1/2)*, 40–54.

Porotova T I (1990). *Kategorija mnozhestvennosti v enisejskikh jazykakh*. Tomsk: Izdatel'stvo tomskogo universiteta.

Shabaev V G (1987). 'Dvojnaja markirovka subekta v sostave finitnykh form ketskogo glagola.' In Morev J A & Denning R F (eds.) *Jazyki narodov SSSR*. Novosibirsk: Novosibirsk State University Press. 100–106.

Vajda E J (1999). *Languages of the World 15. Ket prosodic phonology*. Munich: Lincom.

Vajda E J (2001a). *Yeniseian peoples and languages*. Surrey: Curzon Press.

Vajda E J (2001b). 'The role of position class in Ket verb morphophonology.' *Word 52(3)*, 369–436.

Vajda E J (2002). 'The origin of phonemic tone in Ket.' *Chicago Linguistic Society 37(2)*, 290–305.

Vajda E J (2003). 'Ket verb structure in typological perspective.' In Vajda E & Anderson G D S (eds.) *Studia Yeniseica: typological studies on Yeniseic in honor of Heinrich Werner, Sprachtypologie und Universalienforschung 56(1/2)*, 55–92.

Verner G K (1991). *Sravnitel'naja fonetika enisejskikh jazykov*. Taganrog: Taganrogskij pedagogicheskij institut.

Werner H K (1972). 'Die Tonität der gegenwärtigen Jenissei-Sprachen.' *Zeitschrift für PhonetikSprachwissenschaft und Kommunikationsforschung 25(1–2)*, 111–125.

Werner H K (1995). *Zur Typologie der Jenissej-Sprachen*. Wiesbaden: Harrassowitz.

Werner H K (1997a). *Die ketische Sprache*. Tunguso-Sibirica 3. Wiesbaden: Harrassowitz.

Werner H K (1997b). *Abriß der kottischen Grammatik*. Tunguso-Sibirica 4. Wiesbaden: Harrassowitz.

Werner H K (1997c). *Das Jugische (Sym-Ketische)*. Veröffentlichungen der Societas Uralo-Altaica 50. Wiesbaden: Harrassowitz.

Khakas *See:* Turkic Languages.

Khalaj *See:* Turkic Languages.

Khalkha *See:* Mongolic Languages.

Khasi

G D S Anderson, Salem, OR, USA

The Khasi are a group of Mon-Khmer speakers living predominantly in the Khasi and Jaintia Hills region of Megahlaya state in northeastern India, with a smaller number in Assam, West Bengal, and Manipur states. In some sources, the Khasi have been called Khuchia. The vast majority of Khasi people (ca. 90%) live in India, with a further 10% living across the border in Bangladesh. Khasi [KHI] is the only language of the Mon-Khmer group of Austroasiatic spoken this far to the west.

Traditionally the Khasi are divided into a number of 'dialect' groups, but perhaps it is more sound from a linguistic perspective to speak of a small group of related languages sometimes labeled Khasic or Khasian. These 'dialects' or closely related languages include the following (Grimes, 2000; Parkin, 1991):

(1) Khasic languages
- a. Amwi [AML]
- b. Bhoi
- c. Lyngngam
- d. Pnar (aka Synteng or Jaintia) [PBV]
- e. Khynriam or Cherrapunji/Standard Khasi [KHI]
- f. War

Of these dialects/languages, Lyngngam is most linguistically distant from the standard Khasi dialect, while Pnar is the closest. Amwi is also quite distant from Standard Khasi. Lyngngam may include a linguistically Khasified Garo element (a Tibeto-Burman language), while War and Bhoi may include assimilated Mikir (Tibeto-Burman) elements. The Pnar (Synteng/Jaintia) ruled a kingdom in the region from at least 1500 to 1835, when it was disbanded by the British colonial authorities (Parkin, 1991: 58). Further Khasic varieties include Lakadong and Mynnar. Other local Khasic varieties may (in fact probably do) exist, and there is also considerable microlevel variation.

According to figures from the India Missions Association in 1997, there were just under 1 million total Khasi speakers, including all the above mentioned dialects/languages (the actual estimated figure is 950 000). Khasi is a literary language and a language of media and government in Meghalaya. There are even radio and television broadcasts in the Standard Khasi language. Phonologically, Khasi exhibits some areally and typologically atypical initial clusters, e.g., [bt], [ks], [kt], [ktH], and so on.

(2) *bta* 'wash/besmear face' *ksew* 'dog' *kti* 'hand' *kt*^{h}*áw* 'grandfather'

Syntactically, Khasi is SVO, while other Khasic languages show different basic word orders as well as many other different features.

(3) <u>Khasi</u>
phi-m ʔiithuʔ ya ŋa
you-NEG recognize OBJ I
'don't you recognize me?'
(Rabel, 1961: 61)

Morphosyntactically, Khasi is characterized by use of gender markers and a system of personal verb inflection, albeit within a predominantly isolating structure.

(4) <u>Standard Khasi</u>
u khinnaʔ u-m bam
DET.M boy MASC-NEG eat
'the boy doesn't eat'
(Nagaraja, 1993: 5)

In addition, there is evidence of a now (mainly) covert noun-class system that manifests itself in the form of lexicalized prefixes in noun stems. This system of gender classifiers is highly marked for Austroasiatic. This and the unusual phonology of Khasi set it apart from its sister languages spoken to the east.

In terms of verbal derivational morphology, Khasi, like its sister languages to the east, makes use of a causative prefix consisting of a labial consonant in various allomorphic realizations.

(5)

i. <u>Khasi</u>	<u>Khasi</u>	<u>Khasi</u>
ph-rung 'penetrate'	*ph-láit* 'clear away'	*b-ta* 'wash/besmear face'
< *rung* 'enter'	< *láit* 'be free'	
(Henderson, 1976b: 487)	(Henderson, 1976b: 487)	(Henderson, 1976b: 487)

Negative occurs in the form of either an enclitic to a subject pronoun or gender agreement marker or proclitic to the verb stem, depending of the tense/aspect value of the clause. Note also the presence of fused subject+tense forms in Standard Khasi as well.

(6) <u>Standard Khasi</u>

phi-m ʔiithuʔ ya ŋa	*ŋan ʔm-thoʔ*
you-NEG recognize OBJ I	I.FUT NEG-write
'don't you recognize me?'	'I'm not writing'
(Rabel, 1961: 61)	

Standard Khasi
u khinnaʔ u-m bam
DET.M boy MASC-NEG eat
'the boy doesn't eat'
(Nagaraja, 1993: 5)

In Bhoi, another language of the Khasic subgroup of Mon-Khmer, the negative has a different phonological shape and occurs between the lexical verb and a postposed gender/agreement marker.

(7) Bhoi
u khannaʔ bam re u
DET.M boy eat NEG MASC
'the boy doesn't eat'
(Nagaraja, 1993: 5)

Bound aspectual or tense morphemes are rare in Eastern Austroasiatic. There is, however, a quasi-bound suffixal past tense marker in *-laʔ* and a future in *-diʔ* in Lyngngam, a language of the Khasic subgroup.

(8) Lyngngam

brə kyu	*di-laʔ*	*liŋba*	*laʔtap*
man 3PL	go-PAST	through	forest

'the men went through the forest'
(Nagaraja, 1996: 43)

nə dənni	*nə di-laʔ*	*tu dənni-diʔ*
I go.NPAST	I go-PAST	he go.NPAST-FUT
'I go'	'I went'	'he will go'

(Nagaraja, 1996: 44)

mi binnəŋ-diʔ	*mi baŋ-laʔ*
you eat.NPAST-FUT	you eat-PAST
'you will eat'	'you ate'

(Nagaraja, 1996: 44)

In terms of nominal derivation, Khasi, like virtually all Austroasiatic languages, may derive deverbal nominals through a process of *-n-* infixation. Note that sometimes the derived noun reflects a more archaic phonological form than the verb stem it (historically) derives from, e.g., the preservation of initial *s-* in the word for 'wing' while the corresponding verb stem 'fly' has shifted this to *h-*.

(9) Khasi

shnong	*sner*
'village' < *shong* 'sit, dwell'	'feather, wing' < *her* 'fly'

(Henderson, 1976b: 517–518)

Like many Austroasiatic subgroups (Anderson and Zide, 2002), Khasic languages show irregular correspondences in the free-forms of nouns, while the corresponding 'underlying' (usually CVC) roots are clearly cognate across the subgroup. Note the following forms in this regard:

(10) Irregular Khasic correspondences

Khasi	Lyngngam	Synteng	Amwi	Lakadong	Mynnar	War	gloss
ksew	*ksu:/'su:*	*ksaw* ~*kswa*	*ksiá*	*ksaw*	*ksow*	*ksià*	'dog'
sim	*sim*	*sim*			*ksem*	*ksem*	'bird'
khmat	*kh'mat*	*khmat*	*ma:t*	*ma:t*		*ma:t*	'eye'
khmut	*leo- 'mut*	*khmut*	*mur-koŋ*	*mur-koŋ*		*myrkoŋ*	'nose'

(Fournier, 1974: 86–92)

That elements like *k-/kh-* are historically prefixes in Khasi is attested to by such facts as the following alternations. The original root forms appear as CVC 'combining form' elements in compounds.

(11) Khasi
kti but ***tiipdeŋ*** 'middle finger.' (Rabel, 1961: 44)
khmat but ***matliʔ*** 'white of eye.'
also ***ʔiimat*** 'eye' < see-eye/face (Rabel, 1961: 149)
khnaay 'mouse. rat' but ***naay****saaw* 'small red hill mouse'

Note in this regard also the following alternations:

(12) Khasi
kpa, kmi(e) 'father, mother' (non-vocative) vs.
ʔii paa ʔii mey address term used by children to parents 'da ddy/mommy' (Rabel, 1961: 49)

See also: Austroasiatic Languages; India: Language Situation; Mon-Khmer Languages.

Bibliography

Anderson G D S & Zide N H (2002). 'Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: the bimoraic constraint' In Macken M A (ed.) *Papers from the 10th annual meeting of the Southeast Asian Linguistics Society.* Tempe, AZ: Arizona State University. 55–74. [South East Asian Studies Program, Monograph Series Press.]
Bars Rev E (1973). *Khasi-English dictionary.* Shillong: Don Bosco.
Blah U E (1970). *Khasi self teacher.* Shillong: Chapala Book Stall.
Daladier A (2002). 'Definiteness in Amwi: grammaticalization and syntax.' In Zribi H A & Daladier A (eds.) *Recherches Linguistiques de Vincennes 31.* Paris: Press Universitaires de Vincennes. 61–78.
Fournier A (1974). 'Les Khasi, une population mon-khmer de l'Inde.' *Asie du Sud-Est et Monde Insulindien 5(1),* 79–96.
Gabelentz H C vonder (1858). 'Grammatik und Wörterbuch der Khassia-Sprache.' *Verhandlungen der königlichen Gesellschaft der Wissenschaften zu Leipzig, Philol. -historische Klasse 10,* 1–66.

Grimes B (ed.) (2000). *The Ethnologue* (14th edn.). Dallas: SIL.
Henderson E J A (1965). 'Final -k in Khasi: a secondary phonological pattern.' In Milner G B & Henderson E J A (eds.) *Indo-Pacific linguistic studies.* Amsterdam: North Holland. 459–466.
Henderson E J A (1966). 'Khasi and the l-clusters in seventeenth century Tonkinese.' In Ba Shin U, Boisselier J & Griswold A B (eds.) *Essays offered to G. B. Luce.* (Artibus Asiae, supp. xxiii). Ascona. 139–150.
Henderson E J A (1967). 'Vowel length and vowel quality in Khasi.' *Bulletin of the School of Oriental and African Studies 30,* 564–588.
Henderson E J A (1976a). 'Khasi Initial Clusters.' In Jenner *et al.* (eds.). 523–538.
Henderson E J A (1976b). 'Vestiges of Morphology in Modern Standard Khasi.' In Jenner *et al.* (eds.). 477–522.
Jenner P N, Thompson L C & Starosta S (eds.) (1976). *Austroasiatic studies.* Honolulu: University of Hawaii. [Oceanic Linguistics, Special Publication, No. 13.]
Nagaraja K S (1979). 'Contraction of Khasi nouns in compounds.' *Indian Lingusitics 40,* 18–23.
Nagaraja K S (1984a). 'Compounding in Khasi.' *Bulletin of the Deccan College Research Institute 43,* 79–90.
Nagaraja K S (1984b). 'Reduplication in Khasi.' *Indo-Iranian Journal 27(3),* 189–200.
Nagaraja K S (1985). *Khasi: a descriptive analysis.* Pune, Maharashtra, India: Deccan College.
Nagaraja K S (1993). 'Khasi dialects: a typological consideration.' *Mon-Khmer Studies XXIII,* 1–10.
Nagaraja K S (1996). 'The status of Lyngngam.' *Mon-Khmer Studies 26,* 37–50.
Parkin R (1991). *A guide to Austroasiatic speakers and their languages.* Honolulu: University of Hawaii. [Oceanic Linguistics Special Publication 23.]
Rabel L (1961). *Khasi: a language of Assam.* Baton Rouge: Louisiana State University Press.
Roberts H (1891). *A grammar of the Khassi language.* London: Kegan Paul, Trench, Trübner and Co.
Schmidt Pater W (1904). 'Grundzüge einer Lautlehre der Khasi-Sprache.' *Transactions of Royal Bavarian Academy of Sciences 22(3),* 657–810.
Sharma H S (1999). 'A comparison between Khasi and Manipuri word order.' *Linguistics of the Tibeto-Burman Area 22(1),* 139–148.
Subbarao K V & Temsen G M (2003). 'Wh-question formation in Khasi.' In Kansakar T R & Turin M (eds.) *Themes in Himalayan languages.* Heidelberg: South Asia Institute/ Kathmandu: Tribhuvan University. 197–218.
Weidert A K (1975). *I tkong Amwi. Deskriptive Analyse eines Wardialekts des Khasi.* Wiesbaden: Harrassowitz.

Khmer

M Minegishi, Tokyo University of Foreign Studies, Tokyo, Japan

Khmer (Cambodian)

In the Kingdom of Cambodia, most of the population of 10 716 000 (1998 UN) are considered speakers of Khmer. Its dialectal varieties are spoken by around 1.3 million ethnic Khmer people in the northeastern and eastern provinces of Thailand. The former variety is called Northern Khmer, or sometimes Surin Khmer. Another variety is spoken by more than one million people of the Khmer ethnic group, called Lower Khmer, in southern Vietnam.

Khmer is one of the major languages of Mon-Khmer subgroup of the Austroasiatic language family. It is a typical Mon-Khmer language in that phonemically its native words are either monosyllabic or disyllabic; it has no tonal distinction and is of the isolating type syntactically.

Script and Written Records

Khmer script is one of the oldest scripts in mainland Southeast Asia that originate in South India. Archeology has shown that the communication between Southeast Asia and India dates back to the beginning of the Christian era. Khmer legend, according to the Chinese document, says the local queen married the prince from India, and they became the founders of the Khmer kingdom, which suggests the existence of local matrimonial authority influenced by the Indian civilization, such as Hinduism.

The oldest Khmer inscription in old Khmer dates from 611 C.E., in addition to which there are also undated inscriptions, or presumably older ones written in Sanskrit. The inscriptions spread not only in Cambodia, but also in parts of Thailand and southern Vietnam, which suggests that the ethnic group was formerly more widespread than it is in the present and once exerted a strong cultural influence over the area.

Phonology and Phonetics

A Khmer native word is either monosyllabic or disyllabic. In a disyllabic word, a minor syllable precedes a major syllable. In a minor syllable, the inventory of possible vowels is smaller than in a major syllable. A major syllable is pronounced with stress when preceded by a minor syllable. Using the abbreviations

Grimes B (ed.) (2000). *The Ethnologue* (14th edn.). Dallas: SIL.
Henderson E J A (1965). 'Final -k in Khasi: a secondary phonological pattern.' In Milner G B & Henderson E J A (eds.) *Indo-Pacific linguistic studies.* Amsterdam: North Holland. 459–466.
Henderson E J A (1966). 'Khasi and the l-clusters in seventeenth century Tonkinese.' In Ba Shin U, Boisselier J & Griswold A B (eds.) *Essays offered to G. B. Luce.* (Artibus Asiae, supp. xxiii). Ascona. 139–150.
Henderson E J A (1967). 'Vowel length and vowel quality in Khasi.' *Bulletin of the School of Oriental and African Studies 30,* 564–588.
Henderson E J A (1976a). 'Khasi Initial Clusters.' In Jenner *et al.* (eds.). 523–538.
Henderson E J A (1976b). 'Vestiges of Morphology in Modern Standard Khasi.' In Jenner *et al.* (eds.). 477–522.
Jenner P N, Thompson L C & Starosta S (eds.) (1976). *Austroasiatic studies.* Honolulu: University of Hawaii. [Oceanic Linguistics, Special Publication, No. 13.]
Nagaraja K S (1979). 'Contraction of Khasi nouns in compounds.' *Indian Lingusitics 40,* 18–23.
Nagaraja K S (1984a). 'Compounding in Khasi.' *Bulletin of the Deccan College Research Institute 43,* 79–90.
Nagaraja K S (1984b). 'Reduplication in Khasi.' *Indo-Iranian Journal 27(3),* 189–200.
Nagaraja K S (1985). *Khasi: a descriptive analysis.* Pune, Maharashtra, India: Deccan College.
Nagaraja K S (1993). 'Khasi dialects: a typological consideration.' *Mon-Khmer Studies XXIII,* 1–10.
Nagaraja K S (1996). 'The status of Lyngngam.' *Mon-Khmer Studies 26,* 37–50.
Parkin R (1991). *A guide to Austroasiatic speakers and their languages.* Honolulu: University of Hawaii. [Oceanic Linguistics Special Publication 23.]
Rabel L (1961). *Khasi: a language of Assam.* Baton Rouge: Louisiana State University Press.
Roberts H (1891). *A grammar of the Khassi language.* London: Kegan Paul, Trench, Trübner and Co.
Schmidt Pater W (1904). 'Grundzüge einer Lautlehre der Khasi-Sprache.' *Transactions of Royal Bavarian Academy of Sciences 22(3),* 657–810.
Sharma H S (1999). 'A comparison between Khasi and Manipuri word order.' *Linguistics of the Tibeto-Burman Area 22(1),* 139–148.
Subbarao K V & Temsen G M (2003). 'Wh-question formation in Khasi.' In Kansakar T R & Turin M (eds.) *Themes in Himalayan languages.* Heidelberg: South Asia Institute/ Kathmandu: Tribhuvan University. 197–218.
Weidert A K (1975). *I tkong Amwi. Deskriptive Analyse eines Wardialekts des Khasi.* Wiesbaden: Harrassowitz.

Khmer

M Minegishi, Tokyo University of Foreign Studies, Tokyo, Japan

Khmer (Cambodian)

In the Kingdom of Cambodia, most of the population of 10 716 000 (1998 UN) are considered speakers of Khmer. Its dialectal varieties are spoken by around 1.3 million ethnic Khmer people in the northeastern and eastern provinces of Thailand. The former variety is called Northern Khmer, or sometimes Surin Khmer. Another variety is spoken by more than one million people of the Khmer ethnic group, called Lower Khmer, in southern Vietnam.

Khmer is one of the major languages of Mon-Khmer subgroup of the Austroasiatic language family. It is a typical Mon-Khmer language in that phonemically its native words are either monosyllabic or disyllabic; it has no tonal distinction and is of the isolating type syntactically.

Script and Written Records

Khmer script is one of the oldest scripts in mainland Southeast Asia that originate in South India. Archeology has shown that the communication between Southeast Asia and India dates back to the beginning of the Christian era. Khmer legend, according to the Chinese document, says the local queen married the prince from India, and they became the founders of the Khmer kingdom, which suggests the existence of local matrimonial authority influenced by the Indian civilization, such as Hinduism.

The oldest Khmer inscription in old Khmer dates from 611 C.E., in addition to which there are also undated inscriptions, or presumably older ones written in Sanskrit. The inscriptions spread not only in Cambodia, but also in parts of Thailand and southern Vietnam, which suggests that the ethnic group was formerly more widespread than it is in the present and once exerted a strong cultural influence over the area.

Phonology and Phonetics

A Khmer native word is either monosyllabic or disyllabic. In a disyllabic word, a minor syllable precedes a major syllable. In a minor syllable, the inventory of possible vowels is smaller than in a major syllable. A major syllable is pronounced with stress when preceded by a minor syllable. Using the abbreviations

C for a consonant, V for a vowel, F for a syllable final consonant, 'r' for a liquid and parentheses for an optional element, a minor syllable can be either CVF or C(r)V, where only nasals can appear as F.

Likewise, a Khmer major syllable can be illustrated as C (C) V F. Vowels are either long or short. The long and short contrast is also found among diphthongs. A short vowel occurs only in a checked syllable: it must be followed by a syllable-final consonant, whereas a long vowel can occur both in an open and checked syllable.

Consonants

Consonants in the syllable-initial position are given in Table 1, where the IPA symbols, when necessary, are given in the brackets.

Of these consonants, only nasals /m, n, ɲ ŋ/, unreleased stops /p˺, t˺, c˺, k˺/, a glottal stop /ʔ/, fricatives /v, y, h/ and a liquid /l/ can occur in the word-final position as well. /r/ in the word-final position, once pronounced, is lost except in the Northern Khmer dialect in Thailand. /f/ appears only in loanwords.

The distinction between voiceless unaspirated and voiceless aspirated stops can be found only in the syllable initial or intervocalic position. Voiceless aspirated consonants could be further analyzed as consonant clusters /p/ + /h/, /t/ + /h/, /c/ + /h/, /k/ + /h/, as there are some words that have an infix between the first stop and the second fricative /h/. Characteristic of the Khmer consonants is that Khmer allows a variety of two-consonant combinations in the syllable initial position of a major syllable. Regarding the above aspirated stops as consonant clusters, 84 combinations in total are possible.

Vowels

Modern Khmer has a complicated vowel system. Although several dictionaries have been published for Khmer, there is no consensus as to the vowel phonemic system. As a result, almost every dictionary has its own phonemic transcription. The main reason for the discrepancy is that some assume the existence of resister contrast, i.e., contrast between 'breathy' and 'clear' phonation type, but others do not.

Table 1 Khmer consonants

p	t	c [tɕ]	k	ʔ
ph	th	ch [tɕh]	kh	
b[ɓ]	d[ɗ]			
m	n	ɲ	ŋ	
	r l			
(f) v	s y[j]			h

Table 2 shows the standard Khmer long vowels and diphthongs in phonetic transcription with the transliteration of orthography of Indic origin. The register contrast, which had been regarded as phonemic, has been lost in the dialects in Cambodian territory, and diphthongization has occurred in compensation. See Minegishi (1985) for details of phonetic values of Khmer dialects. According to Wayland and Jongman (2003), the remnant of phonation contrast is observed in the dialect of eastern Thailand.

Note that, "æ, ɨ̄, ə" are transliterations for Khmer original scripts, which do not exist in the ordinary Indic script system. A '1' following the transliteration means that the consonant preceding the vowel symbol is of voiceless group; '2' indicates the preceding consonant of the voiced group. In addition, there are diphthongs /ɯə, ɔə, uə/. In total, there are 12 long vowels and eight diphthongs. This complexity is attributable to the loss of voiceless and voiced contrast in the syllable-initial position for stops and successive divergence of vowels. See 'Historical Phonology' below for details. Table 3 shows the Khmer short vowels and diphthongs.

Historical Phonology

The phonemic reconstruction by Sakamoto based on Khmer inscriptions has established the old Khmer vowel phonemes as */i, e, ɛ, a, aa, (ɯ), ə, u, o, ɔ, ɒ, ɒɒ, iə, uə/.

By the above reconstruction and modern orthography, diachronic changes in their phonological systems can be internally reconstructed as follows.

Formerly, Khmer had a phonemic contrast, e.g., /*kaa/ and /*gaa/ as its orthography shows, where the

Table 2 Khmer long vowels and diphthongs

Phonemic	*Orthography*
iː	ī-2
eː	e-2
e̞ː	e-1
ɛː	æ-2
aɛ	æ-1
ɯː	ɨ̄-2
ə̞ː	ə̄-2
əː	ɨ̄-1
aə	ə̄-1
iə	ā-2
aː	ā-1
uː	ū-2
o̞u	o-2
oː	ū-1
o̞ː	a-2
ɔː	a-1
ao	o-1

difference is between voiceless and voiced consonants. Later, the vowel following the voiced consonant changed its quality; /*kaa/ and /*gèa/, where phonemic contrast between the consonants still existed and the difference in the vowel register, i.e., 'clear' versus 'breathy' phonation type respectively, was irrelevant. Later on, the voiceless and voiced distinction in stops was lost and the difference in the voice quality in turn carried the phonemic contrast; /*kaa/ and /*kèa/. In the present, the voice quality is no longer phonemic; instead vowel articulation is relevant; /kaː/ and /kiə/. The divergence of vowels is well preserved in the standard Khmer around Phnom Penh area, but in the rest of the country, vowels have merged again to simplify the vowel system. As a result, most of the dialects have only /iː, eː/ as long front vowels, etc. Northern Khmer, conversely, has retained most of the vowel contrast as monophthongs.

Tonal Contrast

As is usual in Mon-Khmer languages, Khmer does not have tones. The only exception, however, is the colloquial style of Phnom Penh dialect, which has acquired a tonal contrast, a level tone versus a raising-falling one: the latter is a compensation for the phonemic change /r/ into /h/, and the successive loss of /h/.

Morphology

Khmer, although syntactically of an isolating type, has a large number of derivational prefixes and infixes, which have been fossilized and are no longer productive in word formation. A word may have either a prefix or infix, but not both. Thus, native words are either monosyllabic or disyllabic. Like other Mon-Khmer languages, Khmer has no suffix. It has prefixes and infixes for causativization, specification, nominalization, intransitivization, repetition, etc., and infixes representing instrument, agent, result, object, etc., of an action.

Grammar

Khmer is an isolating language with no inflection in verbs, nor case marking in nouns. As a result, classifying word classes must be done by means of their distribution and class meaning. Noun modifiers follow the noun, verb modifiers follow the verb.

Table 3 Khmer short vowels and diphthongs

i	ɯ	u
e	ə	o
ɛ		ɔʷ
ɛə	a	ɔ

Khmer's main word classes are as follows, although further classification considering the syntactic distribution is possible: nouns, numerals, classifiers, demonstratives, pronouns, verbs, preverbs, adverbs, expressives, conjunctions, and final particles.

Of these, classifiers are few in number and rarely used except for counting persons, animals, or books. There are two demonstratives. Pronouns are a subclass of nouns, most of which are also used as nouns. Along with titles and kinship terms, choice of pronouns shows relative social positions. Verbs can be further classified as active and stative (adjectival) verbs. Preverbs may precede verbs adding modal meaning, such as 'may, must,' etc., to them. Adverbs may follow a verb. Expressives are a subclass of adverbs, describing noises, shapes, movements, emotions, etc. Final particles may be in the sentence-final position to denote the intentions and emotions of the speaker, etc.

The basic word order is Subject+Verb+(Object). A modified noun (head) is followed by a modifier. Nouns and stative verbs can be used as modifiers. Prepositions (or a noun grammaticalized as a preposition) precede a noun.

The typical noun phrase can be described as follows where optional elements are in parentheses.

Noun+(Verb)+(Numerals)+(Classifier)+(Demonstrative)

In case a clause is used for the noun modifier, the word order is as follows.

Noun+(Relative clause marker+Clause)+(Demonstrative)

A verb may be followed by a noun to form a verb phrase. Several verbs, sometimes with a noun inserted in between, form a serial verb construction, or verb serialization without any change in verbal forms, such as V(N)V(N), etc. In a serial verb construction, two or more verbs may be in various semantic relations, such as an action and its direction, an action and its objective, successive actions, an action and its result, an action and its manner, or an action and its means, etc.

Vocabulary

As one of the earliest languages in Southeast Asia that has accepted Indianization, earlier Hinduism, and later Theravada Buddhism, Khmer has borrowed Sanskrit and Pali loans, especially for religion, administrative, and other cultural vocabulary. It also has exerted a huge influence over adjacent Thai, which in turn borrowed a large number of Khmer words, such as honorific vocabulary used for the royal family. As a result of long-term contact with Thai, Khmer also borrowed many words from Thai.

See also: Austroasiatic Languages; Cambodia: Language Situation; Mon-Khmer Languages; Thailand: Language Situation.

Bibliography

Chantharupanth T & Phromjagarin C (1978). *Surin Khmer–English dictionary.* Bangkok: Institute of Indigenous Language Rsearch, Chulalongkorn University Press.

Diffloth G (2003). 'Austroasiatic language.' *Encyclopedia Britannica 2003* (DVD edn.). Encyclopedia Britannica.

Headley R K Jr, Chhor K, Lim L K, Lim H K & Chun C (1977). *Cambodian–English dictionary.* Washington DC: The Catholic University of America Press.

Henderson E J (1952). 'The main features of Cambodian pronunciation.' *Bulletin of the School of Oriental and African Studies 14(1),* 149–174.

Huffman F (1977). *Cambodian reader and glossary (Yale linguistic series).* New Haven and London: Yale University Press.

Huffman F & Proum I (1977). *English–Khmer dictionary.* New Haven, CT: Yale University Press.

Institute bouddhique (1967). *Dictionnaire cambodgien.* Phnom Penh.

Jacob J M (1968). *An introduction to Cambodian.* London: Oxford University Press.

Jacob J M (1974). *A concise Cambodian–English dictionary.* London: Oxford University Press.

Jenner P N (1974a). 'The development of the registers in Standard Khmer.' In Liem N D (ed.) *Southeast Asian linguistic studies I.* Canberra: Australian National University. 47–60.

Jenner P N (1974b). 'Observations of the Surin dialect of Khmer.' In Liem N D (ed.) *Southeast Asian linguistic studies I.* Canberra: Australian National University. 61–73.

Martin M A (1975). 'Le dialect cambodgien parlé à Tatêy, massif des Cardamomes.' *Asie du Sud-est et Monde Inslindien 6(4),* 71–79.

Maspéro G (1915). *Grammaire de la langue khmère (cambodgien).* Paris: Imprimerie Nationale.

Minegishi M (1985). 'Modern Khmer orthography and its dialects.' [in Japanese] *Gengo Kenkyu 88.*

Sakamoto Y (1970a). 'i, ī, ya, yā de khmer ancien.' *Journal of Southeast Asian Studies 7,* 4. 492–503.

Sakamoto Y (1970b). 'e de khmer ancien.' *Journal of Asian and African Studies 3,* 28–43.

Sakamoto Y (1971). 'Sur quelques voyelles de khmer ancien.' *Journal of Asian and African Studies 4,* 55–73.

Sakamoto Y (1974). 'a et ā de khmer ancien.' *Journal of Asian and African Studies 4,* 75–100.

Sakamoto Y (2001). *Cambodian–Japanese dictionary* (3 vols). Tokyo: ILCAA, Tokyo University of Foreign Studies.

Wayland R & Jongman A (2003). 'Acoustic correlates of breathy and clear vowels: the case of Khmer.' *Journal of Phonetics 31(2),* 181–201.

Khoesaan Languages

W H G Haacke, University of Namibia, Windhoek, Namibia

The recognition of Khoisan (Khoesaan) as one of the four language phyla in Africa received initial impetus mainly through the 'Macro-Khoisan Hypothesis' of Joseph Greenberg, first published in 1950. This controversial phylum comprises his 'Click languages,' which formerly were known as 'Hottentot' and 'Bushman' languages, respectively. Dorothea Bleek (1929) had paved the way for their integration into one family by challenging the prevalent view that Hottentot as a Hamitic language had been influenced by Bushman, and by suggesting instead that Nama (as representative of Hottentot) was a Bushman language with Hamitic admixture. Bleek divided the Bushman languages into Northern, Central, and Southern groups. In essence, these divisions are still recognized today, although their validity is open to challenges (see **Table 1**).

The compounded name Khoisan was coined in 1928 by Leonhardt Schultze to signify the somatoracial relatedness of the Hottentots and Bushmen, with *khoi* ('human being,' correctly spelled *khoe*) representing Hottentot, and *san* – the Nama designation for the Bushmen, meaning 'foragers' (correctly spelt *Sãn* or *Saan*) – Bushman.

Greenberg distinguished South African Khoisan (with the major Northern, Central, and Southern branches) as opposed to the East African isolates Sandawe and Hadza (some 70 000 and 400 speakers, respectively). These views are not unanimously accepted, as genetic relatedness between the major linguistic branches cannot be proved satisfactorily. Although scholars remain divided on the issue of genetic relatedness, the term 'Khoesan' or 'Khoisan' (more correctly spelled 'Khoesaan') is now widely used as a term of convenience to denote all non-Bantu and non-Cushitic click languages of Southern and Eastern Africa.

Only some 30 Khoesaan languages still exist today, with the great majority of languages being extinct.

See also: Austroasiatic Languages; Cambodia: Language Situation; Mon-Khmer Languages; Thailand: Language Situation.

Bibliography

Chantharupanth T & Phromjagarin C (1978). *Surin Khmer–English dictionary.* Bangkok: Institute of Indigenous Language Rsearch, Chulalongkorn University Press.

Diffloth G (2003). 'Austroasiatic language.' *Encyclopedia Britannica 2003* (DVD edn.). Encyclopedia Britannica.

Headley R K Jr, Chhor K, Lim L K, Lim H K & Chun C (1977). *Cambodian–English dictionary.* Washington DC: The Catholic University of America Press.

Henderson E J (1952). 'The main features of Cambodian pronunciation.' *Bulletin of the School of Oriental and African Studies 14(1),* 149–174.

Huffman F (1977). *Cambodian reader and glossary (Yale linguistic series).* New Haven and London: Yale University Press.

Huffman F & Proum I (1977). *English–Khmer dictionary.* New Haven, CT: Yale University Press.

Institute bouddhique (1967). *Dictionnaire cambodgien.* Phnom Penh.

Jacob J M (1968). *An introduction to Cambodian.* London: Oxford University Press.

Jacob J M (1974). *A concise Cambodian–English dictionary.* London: Oxford University Press.

Jenner P N (1974a). 'The development of the registers in Standard Khmer.' In Liem N D (ed.) *Southeast Asian linguistic studies I.* Canberra: Australian National University. 47–60.

Jenner P N (1974b). 'Observations of the Surin dialect of Khmer.' In Liem N D (ed.) *Southeast Asian linguistic studies I.* Canberra: Australian National University. 61–73.

Martin M A (1975). 'Le dialect cambodgien parlé à Tatêy, massif des Cardamomes.' *Asie du Sud-est et Monde Inslindien 6(4),* 71–79.

Maspéro G (1915). *Grammaire de la langue khmère (cambodgien).* Paris: Imprimerie Nationale.

Minegishi M (1985). 'Modern Khmer orthography and its dialects.' [in Japanese] *Gengo Kenkyu 88.*

Sakamoto Y (1970a). 'i, ī, ya, yā de khmer ancien.' *Journal of Southeast Asian Studies 7,* 4. 492–503.

Sakamoto Y (1970b). 'e de khmer ancien.' *Journal of Asian and African Studies 3,* 28–43.

Sakamoto Y (1971). 'Sur quelques voyelles de khmer ancien.' *Journal of Asian and African Studies 4,* 55–73.

Sakamoto Y (1974). 'a et ā de khmer ancien.' *Journal of Asian and African Studies 4,* 75–100.

Sakamoto Y (2001). *Cambodian–Japanese dictionary* (3 vols). Tokyo: ILCAA, Tokyo University of Foreign Studies.

Wayland R & Jongman A (2003). 'Acoustic correlates of breathy and clear vowels: the case of Khmer.' *Journal of Phonetics 31(2),* 181–201.

Khoesaan Languages

W H G Haacke, University of Namibia, Windhoek, Namibia

The recognition of Khoisan (Khoesaan) as one of the four language phyla in Africa received initial impetus mainly through the 'Macro-Khoisan Hypothesis' of Joseph Greenberg, first published in 1950. This controversial phylum comprises his 'Click languages,' which formerly were known as 'Hottentot' and 'Bushman' languages, respectively. Dorothea Bleek (1929) had paved the way for their integration into one family by challenging the prevalent view that Hottentot as a Hamitic language had been influenced by Bushman, and by suggesting instead that Nama (as representative of Hottentot) was a Bushman language with Hamitic admixture. Bleek divided the Bushman languages into Northern, Central, and Southern groups. In essence, these divisions are still recognized today, although their validity is open to challenges (see Table 1).

The compounded name Khoisan was coined in 1928 by Leonhardt Schultze to signify the somatoracial relatedness of the Hottentots and Bushmen, with *khoi* ('human being,' correctly spelled *khoe*) representing Hottentot, and *san* – the Nama designation for the Bushmen, meaning 'foragers' (correctly spelt *Sān* or *Saan*) – Bushman.

Greenberg distinguished South African Khoisan (with the major Northern, Central, and Southern branches) as opposed to the East African isolates Sandawe and Hadza (some 70 000 and 400 speakers, respectively). These views are not unanimously accepted, as genetic relatedness between the major linguistic branches cannot be proved satisfactorily. Although scholars remain divided on the issue of genetic relatedness, the term 'Khoesan' or 'Khoisan' (more correctly spelled 'Khoesaan') is now widely used as a term of convenience to denote all non-Bantu and non-Cushitic click languages of Southern and Eastern Africa.

Only some 30 Khoesaan languages still exist today, with the great majority of languages being extinct.

Table 1 Southern African Khoesaan languages

Branch	*Languages* (*† = [virtually] extinct*)
1. Northern (*Ju*)	!Xung (DC), Ju|'hoan, ||Xau||'e, !'O!xung
2. Southern (*!Ui-Taa*)	
2.1 †*!Ui*	†|Xam, †|'Aunî, †‡Khomanî, †||Xegwi
2.2 *Taa*	!Xõõ (DC), †Kakia
Isolate	‡Hõã
3. Central (*Khoe*)	
3.1. Khoekhoe	
3.1.1. Northern/Namibian	Khoekhoegowab DC (= Nama, Damara, Hai||om, ‡Ākhoe)
3.1.2. †Southern/South African	†!Gora ('Korana'), †Xri ('Griqua'), †Cape Khoekhoe (DC)
3.2. Kalahari Khoe	
3.2.1. Western	Khwe, Buga, ||Ani (DC); Naro (DC); ||Gana, |Gui, ‡Haba (DC);
3.2.2. Eastern	Shua, Ts'ixa, Danisi, |Xaise, †Deti; Kua-Tsua (DC)
4.	†Kwadi (Angola)
5.	Sandawe (Tanzania)

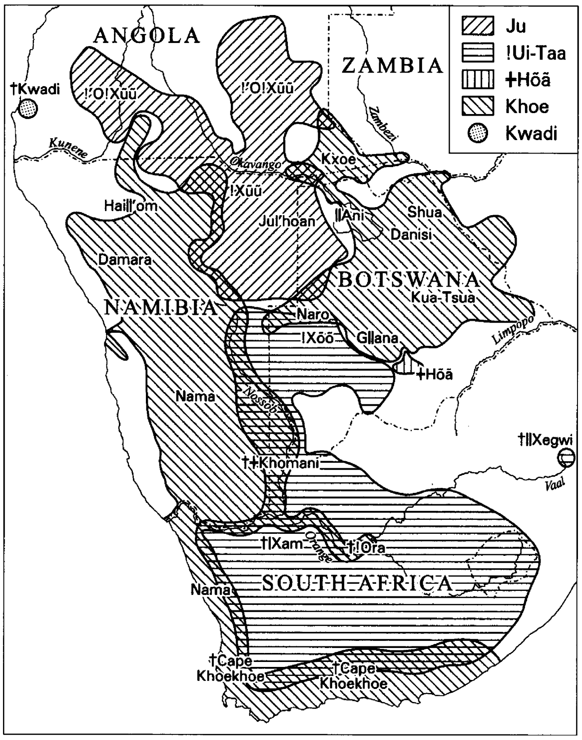


Figure 1 South African Khoesaan languages (precolonial situation). (Note: '†' = language now extinct.) Map reproduced with kind permission of Cambridge University Press from Güldemann T & Vossen R (2000). 'Khoisan.' In Heine B & Nurse, D (eds.). *African languages: an introduction.* Cambridge: Cambridge University Press. 100, map 5.5.

With the possible exception of Khoekhoegowab in Namibia (formerly better known as 'Nama,' and for classificatory purposes briefly referred to as 'Khoekhoe'), virtually all of these languages can be considered to be endangered.

While the Northern (*Ju*) and Southern (*!Ui-Taa*) branches with the isolate ‡Hôâ are spoken by hunter-gatherers (Bushmen/Saan), the languages of the Central branch (*Khoe*) are today spoken by Khoeid (Nama; !Gora/!Ora, Xri, and Cape Khoekhoe being extinct), Saaid (of especially the Kalahari Khoe branch), and Negroid (Damara) peoples. Linguistic and racial classifications of these groups are thus not coextensive. The following classification is largely based on the classifications of Köhler (1989) and of Güldemann and Vossen (2000). The reader is referred to the latter publication for more detailed information on Khoesaan languages. For a classification of !Xung and Jul'hoan dialects (Northern Khoesaan) see Snyman (1997), for Central Khoesaan see Voßen (1997), and for dialects of Khoekhoe(gowab) Haacke *et al.* (1997). Several of the language names below represent dialect clusters (DCs). The iconic classificatory names *Ju*, *!Ui*, *Taa*, and *Khoe* mean 'human being' in their respective branches. The now extinct Kwadi was probably related to Namibian Khoekhoe. Although Sandawe does show evidence of affinity to Central Khoesaan, *i.a.* also in its tonology, recent research by Bonny Sands (1998b) considers Hadza of Tanzania to be an isolate with no satisfactory evidence for a genetic relationship to Khoesaan, despite lexical similarities and the use of clicks (Figure 1).

Before the Bantu diaspora, Khoesaan peoples probably inhabited the entire Southern Africa up to the east coast of South Africa and into southern Angola. Displacement and absorption led to a drastic reduction of territorial domains, which was further aggravated by the arrival of European colonizers in the seventeenth century. Social marginalization still determines the life of all groups considered to be Saan. Khoekhoegowab in Namibia is the only Khoesaan language that is officially recognized for language-planning purposes and is a major subject at university level. Demographic figures are largely based on estimates, totalling over 200 000 for Southern African Khoesaan. Khoekhoe(gowab), with 175 554 speakers (1991 census of Namibia), represents by far the largest Khoesaan speech community and constitutes some 12.5% of the Namibian population.

The most conspicuous phonological characteristic of Khoesaan languages is the use of click consonants (Table 2). Clicks consist of an influx and an efflux phase. The influx (basic click) produces the actual

Table 2 The 20 click variants of Khoekhoegowab

Influx	*Efflux*				
	Glottal stop	*Voiceless velar stop*	*Delayed glottal fricative*	*Voiceless velar affricate*	*Voiced nasalization*
Dental ǀ	[ǀʔ]	ǀg [ǀ]	ǀh [ǀɦ]	ǀkh [ǀx^h]	ǀn [ŋǀ]
Alveolar	! [!ʔ]	!g [!]	!h [!ɦ]	!kh [!x^h]	!n [ŋ!]
Palatal	ǂ [ǂʔ]	ǂg [ǂ]	ǂh [ǂɦ]	ǂkh [ǂx^h]	ǂn [ŋǂ]
Lateral	ǁ [ǁʔ]	ǁg [ǁ]	ǁh [ǁɦ]	ǁkh [ǁx^h]	ǁn [ŋǁ]

Table 3 The six main citation melodies of Khoekhoegowab and their sandhi correlates (as recorded by Eliphas Eiseb)

Citation	*Sandhi*	*Gloss*
!òm̀s	*!òm̏s*	to butt
!ȍḿs	*!ȍḿs*	female genitals, udder
!òm̀s	*!òm̀s*	to force escape from burrow (of: aardvark)
!óm̀s	*!òm̏s*	to coagulate;
		to remove thorn with aid of utensil
!óḿs	*!óm̀s*	fist
!òm̋s	*!òm̀s*	pollard

clicking sound and is produced without pulmonic airstream. There are five influx variants: ⊙ (bilabial), ǀ (dental), ! (alveolar), ǂ (palatal), and ǁ (lateral). Each of these influxes then combines with a specific number of effluxes, depending on the language. This efflux constitutes the resumption of the pulmonic egressive airstream, and its nature depends on the manner of release of the posterior, velaric (and at times glottalic) closure. The bilabial 'kiss-click' ⊙ is manifest only in Southern Khoesaan. The number of effluxes can vary from five in Khoekhoe(gowab) to 16 in !Xõõ, yielding a total of 83 click variants in the latter language, with a total of 117 consonants (cf. Traill, 1985) – a possible world record.

The tonology of Central Khoesaan languages is typologically most akin to those of Southeast Asian languages on account of perturbational (sandhi) processes and the interaction of tonal and segmental phonemes, with depressor consonants triggering tonogenesis (development of contrastive pitch as compensation for the depletion of consonantal contrasts).

According to research thus far, available roots in Khoesaan languages appear to be generally bimoraic. Roots of at least Central Khoesaan languages are disyllabic, with syllable and mora being in isomorphic relation; hence tonal melodies (**Table 3**) consist of a sequence of two register (level) tonemes. The following minimal set illustrates the six main melodies of Khoekhoegowab in citation and sandhi form (Haacke, 1999. The second syllable of *!om* consists of a syllabic nasal [m]. Verbs are quoted here in their infinitival form with the third person fem. sg. pgn-marker *s*). Vowel qualities generally vary between oral and nasalized vowels; in addition, pharyngealized, laryngealized, and breathy vowels or their combinations are found in most non-Khoekhoe languages.

The most distinctive morphological characteristic of Central Khoesaan languages is that they mark nouns for sex gender with postclitic person–gender–number markers, whereas Northern and Southern languages do not. This occurs most consistently in Khoekhoegowab, which marks nouns for person (third, as well as first or second), gender (masculine, feminine, neuter/common), and number (singular, dual, and plural). Non-Central languages have little inflectional morphology. Whereas Non-Central languages have SVO constituent order, Central languages have SOV order in the case of lexically specified NPs.

See also: Botswana: Language Situation; Namibia: Language Situation; Phonetics, Articulatory; Phonetic Classification; South Africa: Language Situation.

Bibliography

Beach D M (1938). *The phonetics of the Hottentot language.* Cambridge: Heffer & Sons.

Bleek D F (1929). *Comparative vocabularies of Bushman languages.* Cambridge: Cambridge University Press.

Güldemann T & Vossen R (2000). 'Khoisan.' In Heine B & Nurse D (eds.) *African languages. An introduction.* Cambridge: Cambridge University Press.

Haacke W H G (1999). 'The tonology of Khoekhoe (Nama/Damara).' In Voßen R (ed.) *Quellen zur Khoisan-Forschung 16.* Cologne: Köppe.

Haacke W H G, Eiseb E & Namaseb L (1997). 'Internal and external relations of Khoekhoe dialects: a preliminary survey.' In Haacke W H G & Elderkin E D (eds.) (1997). 125–209.

Haacke W H G & Elderkin E D (eds.) (1997). *Namibian languages: reports and papers.* Namibian African studies 4. Cologne: Köppe.

Köhler O (1989). *Die Welt der Kxoé-Buschleute.* Berlin: Reimer.

Sands B (1998a). *Research in Khoisan Studies 14: Eastern and Southern African Khoisan.* Cologne: Köppe.

Sands B (1998b). 'The linguistic relationship between Hadza and Khoisan.' In Schladt M (ed.) *Language, identity, and conceptualization among the Khoisan.* Research in Khoisan Studies 15. Cologne: Köppe. 265–283.

Schultze L (1928). *Zur Kenntnis des Körpers der Hottentotten und Buschmänner.* Jena: G. Fischer.
Snyman J W (1975). *Žul'hõasi fonologie en woordeboek.* Cape Town: Balkema.
Snyman J W (1997). 'A preliminary classification of the !Xũũ and Žul'hõasi dialects.' In Haacke W H G & Elderkin E D (eds.). 21–106.
Traill A (1985). *Quellen zur Khoisan Forschung 1: Phonetic and phonological studies of !Xõõ Bushman.* Hamburg: Buske.
Traill A (n.d.) Extinct South African Khoisan Languages (audio-CD with booklet). Johannesburg: University of the Witwatersrand.
Voßen R (1997). Die Khoe-Sprachen. *Ein Beitrag zur Erforschung der Sprachgeschichte Afrikas.* Research in Khoisan Studies 12. Cologne: Köppe.
Westphal E (1971). 'The click languages of southern and eastern Africa.' In Sebeok T (ed.) *Linguistics in sub-Saharan Africa.* The Hague: Mouton. 367–420.

Khoikhoi *See:* Khoesaan Languages.

Khoisan Languages *See:* Khoesaan Languages.

Khotanese

P O Skjærvø, Harvard University, Cambridge, MA, USA

Khotanese, an Eastern Middle Iranian language spoken in the kingdom of Khotan in southwestern Xinjiang, is known from manuscripts on paper and wood found in the Khotan area, as well as in the caves at Dunhuang. The related language Tumshuqese, spoken in Kucha in northwestern Xinjiang, is much less well known. They were written in the southern and northern variants of Brahmi, respectively.

Three stages of Khotanese can been distinguished: Old, Middle, and Late, corresponding to texts from, roughly, the 5th–6th, 7th–8th, and 9th–10th centuries, up to the end of Buddhism in Khotan (ca. 1000).

Written remains consist mainly of Buddhist texts from all three periods, economical and legal documents from the Middle Khotanese period, and letters from the Late Khotanese period.

Khotanese and Tumshuqese contitute the northeastern branch of the Iranian languages, in which Indo-Iran. *ćw, ʒ́w* [tʃw, dʒw] became *ś* [ʃ], *ź* [ʒ]. This feature is only shared by modern-day Wakhi spoken in northeastern Afghanistan.

The phonology of Khotanese is of the Middle Iranian type in which, for instance, *č* [tʃ] and *ǰ* [dʒ], spelled *c* and *j*, had become [ts] and [dz], spelled *tc*, and *js* in the Brahmi alphabet. There are two non-Indic vowel marks, transcribed as *-ä* ([-ə]?) and *-ei*.

Khotanese had retroflex consonants (*ṭ, ṭh, ḍ, ṇ*) in both indigenous words (in which, e.g., *ḍ* < *rd*, *ṇ* < *žn*) and in Indic loanwords. It is not clear whether the aspirated stops, *kh*, etc., were spirants [x], etc., or actually aspirated stops [k^h], etc. It is possible that they were originally spirants, as in other Iranian languages (cf. Khot. *khara-* 'donkey,' Avestan *xara-*, Persian *xar*, etc.), and only later became stops, as suggested by the way Chinese was written in Brahmi in Khotan.

The non-Indic sound *z* was written *ys*. The voiced sibilants *ź* and *ž* were originally not distinguished in writing from *ś* and *ṣ*, but later various strategies for distinguishing them were invented (written double = unvoiced, single = voiced; a subscript curved line (transliterated as ') to indicate the voiced pair; e.g., *śärä* and *śśärä* [*śərə*] 'good,' but *śätä* [*źədə*], *śe* [*źe*] 'second'); the line also probably indicated rhotasized vowels (*nei* [*nei*^{r}] 'nectar' < **nāžā*). There was a single-flap *r* and a trilled *r*.

While intervocalic voiced stops had already been lost in Old Khotanese, intervocalic *k* and *t* still remained in the oldest texts (phonemically *g* and *d*), and final *-i*, *-ä*, and *-u* were still distinct. The development from Old via Middle to Late Khotanese in the

Schultze L (1928). *Zur Kenntnis des Körpers der Hottentotten und Buschmänner.* Jena: G. Fischer.

Snyman J W (1975). *Žuǀ'hõasi fonologie en woordeboek.* Cape Town: Balkema.

Snyman J W (1997). 'A preliminary classification of the !Xų̃ų̃ and Žuǀ'hõasi dialects.' In Haacke W H G & Elderkin E D (eds.). 21–106.

Traill A (1985). *Quellen zur Khoisan Forschung 1: Phonetic and phonological studies of !Xõõ Bushman.* Hamburg: Buske.

Traill A (n.d.) Extinct South African Khoisan Languages (audio-CD with booklet). Johannesburg: University of the Witwatersrand.

Voßen R (1997). Die Khoe-Sprachen. *Ein Beitrag zur Erforschung der Sprachgeschichte Afrikas.* Research in Khoisan Studies 12. Cologne: Köppe.

Westphal E (1971). 'The click languages of southern and eastern Africa.' In Sebeok T (ed.) *Linguistics in sub-Saharan Africa.* The Hague: Mouton. 367–420.

Khoikhoi *See:* Khoesaan Languages.

Khoisan Languages *See:* Khoesaan Languages.

Khotanese

P O Skjærvø, Harvard University, Cambridge, MA, USA

Khotanese, an Eastern Middle Iranian language spoken in the kingdom of Khotan in southwestern Xinjiang, is known from manuscripts on paper and wood found in the Khotan area, as well as in the caves at Dunhuang. The related language Tumshuqese, spoken in Kucha in northwestern Xinjiang, is much less well known. They were written in the southern and northern variants of Brahmi, respectively.

Three stages of Khotanese can been distinguished: Old, Middle, and Late, corresponding to texts from, roughly, the 5th–6th, 7th–8th, and 9th–10th centuries, up to the end of Buddhism in Khotan (ca. 1000).

Written remains consist mainly of Buddhist texts from all three periods, economical and legal documents from the Middle Khotanese period, and letters from the Late Khotanese period.

Khotanese and Tumshuqese contitute the northeastern branch of the Iranian languages, in which Indo-Iran. *ćw, ʹjw* [tʃw, dʒw] became *ś* [ʃ], *ź* [ʒ]. This feature is only shared by modern-day Wakhi spoken in northeastern Afghanistan.

The phonology of Khotanese is of the Middle Iranian type in which, for instance, *č* [tʃ] and *ǰ* [dʒ], spelled *c* and *j*, had become [ts] and [dz], spelled *tc*, and *js* in the Brahmi alphabet. There are two non-Indic vowel marks, transcribed as *-ä* ([-ə]?) and *-ei.*

Khotanese had retroflex consonants (*ṭ, ṭh, ḍ, ṇ*) in both indigenous words (in which, e.g., *ḍ* < *rd*, *ṇ* < *žn*) and in Indic loanwords. It is not clear whether the aspirated stops, *kh*, etc., were spirants [x], etc., or actually aspirated stops [k^h], etc. It is possible that they were originally spirants, as in other Iranian languages (cf. Khot. *khara-* 'donkey,' Avestan *xara-*, Persian *xar*, etc.), and only later became stops, as suggested by the way Chinese was written in Brahmi in Khotan.

The non-Indic sound *z* was written *ys*. The voiced sibilants *ź* and *ž* were originally not distinguished in writing from *ś* and *ṣ*, but later various strategies for distinguishing them were invented (written double = unvoiced, single = voiced; a subscript curved line (transliterated as ') to indicate the voiced pair; e.g., *śärä* and *śśärä* [*śərə*] 'good,' but *śätä* [*źədə*], *śe* [*źe*] 'second'); the line also probably indicated rhotasized vowels (*nei* [*nei*^{r}] 'nectar' < **näžä*). There was a single-flap *ɾ* and a trilled *r*.

While intervocalic voiced stops had already been lost in Old Khotanese, intervocalic *k* and *t* still remained in the oldest texts (phonemically *g* and *d*), and final *-i*, *-ä*, and *-u* were still distinct. The development from Old via Middle to Late Khotanese in the

main involved the loss of distinction between final short vowels and their loss after nasals and in some other positions (e.g., Okhot. *aysu* 'I' > MKhot. *aysä*, LKhot. *a*; OKhot. NOM/ACC *suhāvatānä/u* 'utilities' > MKhot. *suhāvāṃ* > LKhot. *suhāvau*); and the merger of *uṃ* and *āṃ* (e.g., OKhot. GEN-DAT PL *rruṃdänu* 'the kings' > MKhot. *rrāṃdāṃ* > LKhot. *rraudau* [approx. [rɔːdɔː]).

The original six-case system is reduced in Middle Khotanese by the merger of the nominative and accusative (except in pronouns), and further reduced in Late Khotanese to a three- or two-case system. There are three genders, the neuter being of limited use.

The verbal system is of the Eastern Middle Iranian type. There are two stems, present and past (e.g., PRES *bar-*, PAST *buḍa-*; PRES *häm-* 'become,' PAST *häm-äta-*). Khotanese has preserved the Old Iranian moods (indicative, imperative, subjunctive, optative, injunctive), as well as active and middle. The past of intransitive verbs is of the common Iranian type (*buḍ-ä mä* 'carry.PAST.INTRANS-SING.MASCCOP.PRES.1ST.SING' = 'I was carried' > 'I rode'), while that of transitive verbs is based on an active past participle plus copula (*buḍ-e* < **bṛta-āh* 'carry.PAST.TRANS-SING.MASC [COP. 3RD.SING = Ø]' = 'he carried,' *buḍ-ātä* 'she carried'; *buḍ-aimä* < **bṛta-āh ahmi* [COP.PRES.1ST. SING] 'I [MASC] carried,' *buḍāndä* 'they carried,' *buḍ-āndä mä* [COP.PRES.1ST.PL] 'we carried').

Perfect/pluperfect and modal forms are formed from the past tense (e.g., perfect: *nei* [<*ne* + *ī*] *hvat-e śtä balysä* 'NEG-EMPH.PART speak.PAST.TRANS.-SING. MASC COP.3RD.SING Buddha-SING.NOM' = 'the Buddha has not at all said'; pluperfect: *cīyä rr-e bäysānd-ä vät-ä* 'when king-SING.NOM awaken.PAST. INTRANS-MASC COP.PAST.INTR-3RD.SING-MASC' = 'when the king had awakened'; pluperfect optative: *ka nä va ysār-u gyast-a balys-a dāt-u hvat-āndä v-ī-ro ne gāvu vamas-īro* 'if they.ENCL.OBL PARTICLE 1000. SING.NEUT lord.PL.NOM-ACC buddha.PL.-NOM-ACC law. SING.ACC speak.PAST.TRANS-PL.MASC/FEM COP-OPT-3RD.PL NEG at-all understand.OPT.3RD.PL' = 'even even if a thousand lord buddhas had spoken the law to them, they would not at all understand').

The 'potentialis' is formed with a past participle with the ending-*u* (SING ACC NEUT) and the verbs *yan-* 'to do' (active) and *häm-*'become' (passive) and expresses possibility and completion of action (e.g., *ne hvat-a häm-āre* 'NEG speak.PAST.PART-PLUR.MASC become.PRES.-3RD.PL' = 'they cannot be said/ expressed,' *ne hvat-u yan-īmä* 'NEG speak.PAST.PART. NEUT do.PRES-1ST.SING' = 'I cannot say (it),' *cīyä hvat-u yuḍ-āndä* 'when speak.PAST.PART.NEUT do.PAST.TRANS-3RD.PL = 'when they had spoken').

Tumshuqese also has the old augmented imperfect (e.g., *a-ch-i* 'PAST-go-3RD.SING' = 'he went, he has gone'), though the augment may be added only to monosyllabic forms (cf. *bar-i* 'he carried').

The lexicon contains numerous borrowings from Indic, both Middle Indic and Classical Sanskrit. In the Middle and Late Khotanese periods, we also find a small number of Chinese and Tibetan words.

See also: Iran: Language Situation.

Bibliography

Bailey H W (1979). *Dictionary of Khotan Saka.* Cambridge: Cambridge University Press.

Emmerick R E (1968). *Saka grammatical studies.* London: Oxford University Press.

Emmerick R E (1989). 'Khotanese and Tumshuqese.' In Schmitt R (ed.) *Compendium Linguarum Iranicarum.* Wiesbaden: Reichert. 204–229.

Skjærvø P O (2002). *Khotanese manuscripts from Chinese Turkestan in The British Library: a complete catalogue with texts and translations.* Sims-Williams U Contributor. London: The British Library.

Kiel/Kiliaan, Cornelis van (1530–1607)

L De Grauwe, Ghent University, Ghent, Belgium

Cornelis Abts alias van Kiele (he signed as C. Kiel or Cornelis van Kiel), the first great lexicographer of what is now the Dutch language, was born in Duffel (now in the Belgian province of Antwerp, then part of the Duchy of Brabant) not before 1530. His name was latinized into Kilianus (hence Modern Dutch Kiliaan). He may have studied law and the three Biblical languages at Louvain University. The book-keeping diary of the Antwerp printer and publisher Christophe Plantin mentions his name in February 1558 as a compositor and in 1565 as a corrector, which he was to remain until his death on Easter Sunday, April 15, 1607.

Apart from this work, Plantin charged him with translations into Dutch of grammars and dictionaries, e.g., Robert Estiennes *Dictionarium Latinogallicum* (translated 1564; edition now lost); the Dutch section in the anonymous *Dictionarium Tetraglotton*

main involved the loss of distinction between final short vowels and their loss after nasals and in some other positions (e.g., Okhot. *aysu* 'I' > MKhot. *aysä*, LKhot. *a*; OKhot. NOM/ACC *suhāvatānä/u* 'utilities' > MKhot. *suhāvāṃ* > LKhot. *suhāvau*); and the merger of *uṃ* and *āṃ* (e.g., OKhot. GEN-DAT PL *rruṃdänu* 'the kings' > MKhot. *rrāṃdāṃ* > LKhot. *rraudau* [approx. [rɔːdɔː]).

The original six-case system is reduced in Middle Khotanese by the merger of the nominative and accusative (except in pronouns), and further reduced in Late Khotanese to a three- or two-case system. There are three genders, the neuter being of limited use.

The verbal system is of the Eastern Middle Iranian type. There are two stems, present and past (e.g., PRES *bar-*, PAST *buḍa-*; PRES *häm-* 'become,' PAST *häm-äta-*). Khotanese has preserved the Old Iranian moods (indicative, imperative, subjunctive, optative, injunctive), as well as active and middle. The past of intransitive verbs is of the common Iranian type (*buḍ-ä mä* 'carry.PAST.INTRANS-SING.MASCCOP.PRES.1ST.SING' = 'I was carried' > 'I rode'), while that of transitive verbs is based on an active past participle plus copula (*buḍ-e* < **bṛta-āh* 'carry.PAST.TRANS-SING.MASC [COP. 3RD.SING = Ø]' = 'he carried,' *buḍ-ātä* 'she carried'; *buḍ-aimä* < **bṛta-āh ahmi* [COP.PRES.1ST. SING] 'I [MASC] carried,' *buḍāndä* 'they carried,' *buḍ-āndä mä* [COP.PRES.1ST.PL] 'we carried').

Perfect/pluperfect and modal forms are formed from the past tense (e.g., perfect: *nei* [<*ne* + *ī*] *hvat-e śtä balysä* 'NEG-EMPH.PART speak.PAST.TRANS.-SING. MASC COP.3RD.SING Buddha-SING.NOM' = 'the Buddha has not at all said'; pluperfect: *cīyä rr-e bäysānd-ä vät-ä* 'when king-SING.NOM awaken.PAST. INTRANS-MASC COP.PAST.INTR-3RD.SING-MASC' = 'when the king had awakened'; pluperfect optative: *ka nä va ysār-u gyast-a balys-a dāt-u hvat-āndä v-ī-ro ne gāvu vamas-īro* 'if they.ENCL.OBL PARTICLE 1000. SING.NEUT lord.PL.NOM-ACC buddha.PL.-NOM-ACC law. SING.ACC speak.PAST.TRANS-PL.MASC/FEM COP-OPT-3RD.PL NEG at-all understand.OPT.3RD.PL' = 'even even if a thousand lord buddhas had spoken the law to them, they would not at all understand').

The 'potentialis' is formed with a past participle with the ending-*u* (SING ACC NEUT) and the verbs *yan-* 'to do' (active) and *häm-*'become' (passive) and expresses possibility and completion of action (e.g., *ne hvat-a häm-āre* 'NEG speak.PAST.PART-PLUR.MASC become.PRES.-3RD.PL' = 'they cannot be said/ expressed,' *ne hvat-u yan-īmä* 'NEG speak.PAST.PART. NEUT do.PRES-1ST.SING' = 'I cannot say (it),' *cīyä hvat-u yuḍ-āndä* 'when speak.PAST.PART.NEUT do.PAST.TRANS-3RD.PL = 'when they had spoken').

Tumshuqese also has the old augmented imperfect (e.g., *a-ch-i* 'PAST-go-3RD.SING' = 'he went, he has gone'), though the augment may be added only to monosyllabic forms (cf. *bar-i* 'he carried').

The lexicon contains numerous borrowings from Indic, both Middle Indic and Classical Sanskrit. In the Middle and Late Khotanese periods, we also find a small number of Chinese and Tibetan words.

See also: Iran: Language Situation.

Bibliography

Bailey H W (1979). *Dictionary of Khotan Saka*. Cambridge: Cambridge University Press.

Emmerick R E (1968). *Saka grammatical studies*. London: Oxford University Press.

Emmerick R E (1989). 'Khotanese and Tumshuqese.' In Schmitt R (ed.) *Compendium Linguarum Iranicarum*. Wiesbaden: Reichert. 204–229.

Skjærvø P O (2002). *Khotanese manuscripts from Chinese Turkestan in The British Library: a complete catalogue with texts and translations*. Sims-Williams U Contributor. London: The British Library.

Kiel/Kiliaan, Cornelis van (1530–1607)

L De Grauwe, Ghent University, Ghent, Belgium

Cornelis Abts alias van Kiele (he signed as C. Kiel or Cornelis van Kiel), the first great lexicographer of what is now the Dutch language, was born in Duffel (now in the Belgian province of Antwerp, then part of the Duchy of Brabant) not before 1530. His name was latinized into Kilianus (hence Modern Dutch Kiliaan). He may have studied law and the three Biblical languages at Louvain University. The bookkeeping diary of the Antwerp printer and publisher Christophe Plantin mentions his name in February 1558 as a compositor and in 1565 as a corrector, which he was to remain until his death on Easter Sunday, April 15, 1607.

Apart from this work, Plantin charged him with translations into Dutch of grammars and dictionaries, e.g., Robert Estiennes *Dictionarium Latinogallicum* (translated 1564; edition now lost); the Dutch section in the anonymous *Dictionarium Tetraglotton*

(Latin–Greek–French–Dutch) edited by Plantin in 1562 was probably also Kiliaan's work. At Plantin's instigation, Kiliaan began to apply himself to a concise *Dictionarium Teutonico-Latinum* (edited in 1574, 240 pp.), starting from his own 'vernacular' mother tongue and comparing the words with their German and French equivalents *(Dictiones Latinè interpretatas, seduloque Germanicis & Gallicis collatas)*. Here at least, *Teutonicus* means 'Dutch,' not 'German' as in some other dictionaries and many other sources of the period – Dutch and German still being regarded as closely related varieties in a single, if internally differentiated, continuum of regional, continental West Germanic '*Schreibsprachen*' (written language variants) that extended from the North and Baltic Seas to the Alps. Kiliaan's lexical basis is his own Brabant (Southern Central Dutch) 'language,' of which he was rather proud, not least because it was spoken in an important region that included cities such as Brussels, Antwerp and Louvain (Leuven). His innovative dictionary, of which only two copies survive, laid the foundation for the modern scientific lexicography of Dutch and indeed for that of all the Western European languages. This is even more true of the two subsequent, much expanded editions. The first of these was the *Dictionarium Teutonico–Latinum* (1588; 765 + 3 pp.), in which Kiliaan included not only additional etymological information but also English, Saxon (= Lower German, Niederdeutsch), Spanish, Italian, and Greek translations as well as new lexical entries from all the Dutch-speaking provinces of the historical Netherlands before their *de facto* separation in 1585 (now 'Holland' and 'Flanders,' the latter constituting northern Belgium). Regional usage is marked by abbreviations such as *fland., holl., fris.* (i.e., used in Flanders, Holland, Friesland); *sic.* stands for *Sicambri,* denoting the people in Gelderland, Kleve, and Jülich, border regions speaking a dialect closely reated to Dutch that now belong entirely or partly to Germany. The third edition, entitled *Etymologicum teutonicae linguae* (1599), is yet further enlarged and in its title reflects the increasing attention paid to etymology. Despite his great merits, Kiliaan remained a child of his time: many of his etymologies are prescientific, and he was a compiler in the literal sense – i.e., a plunderer. In all three editions of his greatest work he tapped into a wide variety of earlier sources, often uncritically, though in the last two editions he did cite them explicitly in an attempt to derive *auctoritas* for himself. Other works by Kiliaan are translations of geographical and historical descriptions and some Latin poems, some of them intended to accompany engravings; from 1574 onward, he also worked as a much-sought-after translator for the Antwerp and Brabant authorities.

See also: Dutch; Etymology; Dutch and Flemish Lexicography.

Bibliography

Claes F (1984). 'Lingua teutonica bij Kiliaan.' *Tijdschrift voor Nederlands en Afrikaans 2,* 157–169.

Kiel C (1574). *Dictionarium Teutonico-Latinum,* introduced by F Claes [Repr. Hildesheim, New York: Olms, 1975].

Kiel C (1599). *Etymologicvm Tevtonicae lingvae sive dictionarivm Tevtonico–Latinvm, praecipvas Tevtonicae lingvae dictiones et phrases Latinè interpretatas* [Repr. Amsterdam: Hakkert, 1972].

Van den Branden L, Cockx-Indestege E & Sillis F (1978). *Bio-bibliografie van Cornelis Kiliaan.* Nieuwkoop: B. de Graaf.

Kilwardby, Robert (ca. 1215–1279)

E Bell Canon, Social Circle, Athens, GA, USA

Robert Kilwardby was born in Leicestershire, England ca. 1215. Beginning in 1231, Kilwardby was a student at the University of Paris, which was considered to be one of the greatest centers of learning in the medieval world. Later, Kilwardby was a master at the University of Paris at a time when debate over Aristotelian philosophy began to divide the church. He was a distinguished teacher of grammar and logic. "In grammar he expounded *Priscianus minor, De accentu* and the *Barbarismus Donati,* covering syntax, prosody, solecism and figurative speech." (Lewry, p. xiv). As a teacher, he paid particular attention to the use of syllogisms in the teaching of logic.

His studies took him from Paris to Oxford, where he ultimately became a master in theology in 1256. He joined the Dominican order as a friar. The Oxford Dominicans divided their studies into two branches: theological studies and philosophical studies. Kilwardby devoted himself to theological subjects

(Latin–Greek–French–Dutch) edited by Plantin in 1562 was probably also Kiliaan's work. At Plantin's instigation, Kiliaan began to apply himself to a concise *Dictionarium Teutonico-Latinum* (edited in 1574, 240 pp.), starting from his own 'vernacular' mother tongue and comparing the words with their German and French equivalents *(Dictiones Latinè interpretatas, seduloque Germanicis & Gallicis collatas)*. Here at least, *Teutonicus* means 'Dutch,' not 'German' as in some other dictionaries and many other sources of the period – Dutch and German still being regarded as closely related varieties in a single, if internally differentiated, continuum of regional, continental West Germanic '*Schreibsprachen*' (written language variants) that extended from the North and Baltic Seas to the Alps. Kiliaan's lexical basis is his own Brabant (Southern Central Dutch) 'language,' of which he was rather proud, not least because it was spoken in an important region that included cities such as Brussels, Antwerp and Louvain (Leuven). His innovative dictionary, of which only two copies survive, laid the foundation for the modern scientific lexicography of Dutch and indeed for that of all the Western European languages. This is even more true of the two subsequent, much expanded editions. The first of these was the *Dictionarium Teutonico–Latinum* (1588; 765 + 3 pp.), in which Kiliaan included not only additional etymological information but also English, Saxon (= Lower German, Niederdeutsch), Spanish, Italian, and Greek translations as well as new lexical entries from all the Dutch-speaking provinces of the historical Netherlands before their *de facto* separation in 1585 (now 'Holland' and 'Flanders,' the latter constituting northern Belgium). Regional usage is marked by abbreviations such as *fland., holl., fris.* (i.e., used in Flanders, Holland, Friesland); *sic.* stands for *Sicambri,* denoting the people in Gelderland, Kleve, and Jülich, border regions speaking a dialect closely reated to Dutch that now belong entirely or partly to Germany. The third edition, entitled *Etymologicum teutonicae linguae* (1599), is yet further enlarged and in its title reflects the increasing attention paid to etymology. Despite his great merits, Kiliaan remained a child of his time: many of his etymologies are prescientific, and he was a compiler in the literal sense – i.e., a plunderer. In all three editions of his greatest work he tapped into a wide variety of earlier sources, often uncritically, though in the last two editions he did cite them explicitly in an attempt to derive *auctoritas* for himself. Other works by Kiliaan are translations of geographical and historical descriptions and some Latin poems, some of them intended to accompany engravings; from 1574 onward, he also worked as a much-sought-after translator for the Antwerp and Brabant authorities.

See also: Dutch; Etymology; Dutch and Flemish Lexicography.

Bibliography

Claes F (1984). 'Lingua teutonica bij Kiliaan.' *Tijdschrift voor Nederlands en Afrikaans 2,* 157–169.

Kiel C (1574). *Dictionarium Teutonico-Latinum,* introduced by F Claes [Repr. Hildesheim, New York: Olms, 1975].

Kiel C (1599). *Etymologicvm Tevtonicae lingvae sive dictionarivm Tevtonico–Latinvm, praecipvas Tevtonicae lingvae dictiones et phrases Latinè interpretatas* [Repr. Amsterdam: Hakkert, 1972].

Van den Branden L, Cockx-Indestege E & Sillis F (1978). *Bio-bibliografie van Cornelis Kiliaan.* Nieuwkoop: B. de Graaf.

Kilwardby, Robert (ca. 1215–1279)

E Bell Canon, Social Circle, Athens, GA, USA

Robert Kilwardby was born in Leicestershire, England ca. 1215. Beginning in 1231, Kilwardby was a student at the University of Paris, which was considered to be one of the greatest centers of learning in the medieval world. Later, Kilwardby was a master at the University of Paris at a time when debate over Aristotelian philosophy began to divide the church. He was a distinguished teacher of grammar and logic. "In grammar he expounded *Priscianus minor, De accentu* and the *Barbarismus Donati,* covering syntax, prosody, solecism and figurative speech." (Lewry, p. xiv). As a teacher, he paid particular attention to the use of syllogisms in the teaching of logic.

His studies took him from Paris to Oxford, where he ultimately became a master in theology in 1256. He joined the Dominican order as a friar. The Oxford Dominicans divided their studies into two branches: theological studies and philosophical studies. Kilwardby devoted himself to theological subjects

almost completely (*De ortu scientiarum*, a basic summary of human science, being the great exception). In order to simplify the research process, he created a detailed set of indices, both alphabetical and analytical, and a set of chapter summaries of the works of Augustine.

In 1261, Kilwardby was elected Provincial Prior of his order, a post he held until 1272. On October 11, 1272, he was appointed Archbishop of Canterbury by Pope Gregory X. He was noted for being charitable to both the poor and the members of his order, giving the necessary money to establish housing for the Blackfriars in London.

In March of 1277, Kilwardby visited Oxford and issued a condemnation of 30 propositions in grammar, logic, and natural science taught there at the time. As a former teacher of logic and grammar, he judged the propositions to be erroneous in nature. The fault, as Kilwardby saw it, was in both the Averroist nature of the propositions and the method of teaching them. Kilwardby was on the Augustinian end of the logic dichotomy that had been seen earlier at the University of Paris. It was not his intention to denounce the propositions as heretical, only to prevent the teaching of them. Kilwardby's condemnations came just days after Etienne Tempier, Bishop of Paris, condemned 219 propositions dealing with Averroist philosophy, advancing the more Christian, Augustinian version of logic taught in the universities.

Pope Nicholas III appointed Kilwardby Cardinal-Bishop of Porto and Santa Rufina on April 4, 1278. Kilwardby was accused of having left Canterbury with church registers and a considerable amount of money. He died in Viterbo, Italy on September 11, 1279. Soon thereafter, the suspicion arose that he had been poisoned, since the money and records were never recovered.

Other works by Kilwardby include *De Tempore*, an inquiry into the nature of time in which the author attempts to reconcile differences between Aristotelian philosophy and Augustinian Christianity, with regard to time and eternity; and *De Universali*.

See also: Aristotle and Linguistics; Priscianus Caesariensis (d. ca. 530).

Bibliography

Hook W F (1865). *Lives of the archbishops of Canterbury* (vol. 3). London: Richard Bentley. 304–326.

Kilwardby R (1987). *On time and imagination, part 1.* Lewry P O (ed.). Oxford: Oxford University Press.

Kilwarby R (1993). *On time and imagination, part 2.* Brodie A (ed.). Oxford: Oxford University Press.

Knight K (1910/2004). *The Catholic Encyclopedia.* In Robert Kilwardby (ed.). Robert Appleton Company.

Weisheipl J A (1967). *New Catholic Encyclopedia.* In Robert Kilwardby (ed.), 19 vols. The Catholic University of America. New York: McGraw-Hill.

Relevant Website

http://www.newadvent.org/ – *The Catholic Encyclopedia.*

Kimatuumbi: Phonology

D Odden, Ohio State University, Columbus, OH, USA

Kimatuumbi is a Bantu language spoken by approximately 60 000 people in Kilwa District of Tanzania. It is a part of the Rufiji-Ruvuma subgroup in the Nurse (1988) classification, designated P13 in Guthrie (1971). General characteristics of the Rufiji-Ruvuma languages are discussed in Odden (2003). As is typically the case in Bantu languages, most phonological processes in the language pertain to prosody – tone, vowel length, and syllable structure – with a bit of segmental phonology regarding vowel harmony and nasal + consonant sequences.

Aspects of Kimatuumbi phonology have been discussed in various publications, for instance Cowper and Rice (1987), Hayes (1990), Odden (1987, 1990), Truckenbrodt (1999), focusing particularly on theoretical consequences of certain phrasal sandhi rules of the language, as well as Rosenthall (1994) discussing vowel sequences. Odden (1996) provides a comprehensive overview of the morphology and phonology of the language.

Segmental Inventory

The 'core' segmental phonemes of the language are as follows:

(1)	p	t	č	k	i		u
	b			g	ɪ		ʊ
	m	n	ñ	ŋ	ɛ		ɔ
	w	l	y			a	

almost completely (*De ortu scientiarum,* a basic summary of human science, being the great exception). In order to simplify the research process, he created a detailed set of indices, both alphabetical and analytical, and a set of chapter summaries of the works of Augustine.

In 1261, Kilwardby was elected Provincial Prior of his order, a post he held until 1272. On October 11, 1272, he was appointed Archbishop of Canterbury by Pope Gregory X. He was noted for being charitable to both the poor and the members of his order, giving the necessary money to establish housing for the Blackfriars in London.

In March of 1277, Kilwardby visited Oxford and issued a condemnation of 30 propositions in grammar, logic, and natural science taught there at the time. As a former teacher of logic and grammar, he judged the propositions to be erroneous in nature. The fault, as Kilwardby saw it, was in both the Averroist nature of the propositions and the method of teaching them. Kilwardby was on the Augustinian end of the logic dichotomy that had been seen earlier at the University of Paris. It was not his intention to denounce the propositions as heretical, only to prevent the teaching of them. Kilwardby's condemnations came just days after Etienne Tempier, Bishop of Paris, condemned 219 propositions dealing with Averroist philosophy, advancing the more Christian, Augustinian version of logic taught in the universities.

Pope Nicholas III appointed Kilwardby Cardinal-Bishop of Porto and Santa Rufina on April 4, 1278. Kilwardby was accused of having left Canterbury with church registers and a considerable amount of money. He died in Viterbo, Italy on September 11, 1279. Soon thereafter, the suspicion arose that he had been poisoned, since the money and records were never recovered.

Other works by Kilwardby include *De Tempore,* an inquiry into the nature of time in which the author attempts to reconcile differences between Aristotelian philosophy and Augustinian Christianity, with regard to time and eternity; and *De Universali.*

See also: Aristotle and Linguistics; Priscianus Caesariensis (d. ca. 530).

Bibliography

Hook W F (1865). *Lives of the archbishops of Canterbury* (vol. 3). London: Richard Bentley. 304–326.
Kilwardby R (1987). *On time and imagination, part 1.* Lewry P O (ed.). Oxford: Oxford University Press.
Kilwarby R (1993). *On time and imagination, part 2.* Brodie A (ed.). Oxford: Oxford University Press.
Knight K (1910/2004). *The Catholic Encyclopedia.* In Robert Kilwardby (ed.). Robert Appleton Company.
Weisheipl J A (1967). *New Catholic Encyclopedia.* In Robert Kilwardby (ed.), 19 vols. The Catholic University of America. New York: McGraw-Hill.

Relevant Website

http://www.newadvent.org/ – *The Catholic Encyclopedia.*

Kimatuumbi: Phonology

D Odden, Ohio State University, Columbus, OH, USA

Kimatuumbi is a Bantu language spoken by approximately 60 000 people in Kilwa District of Tanzania. It is a part of the Rufiji-Ruvuma subgroup in the Nurse (1988) classification, designated P13 in Guthrie (1971). General characteristics of the Rufiji-Ruvuma languages are discussed in Odden (2003). As is typically the case in Bantu languages, most phonological processes in the language pertain to prosody – tone, vowel length, and syllable structure – with a bit of segmental phonology regarding vowel harmony and nasal + consonant sequences.

Aspects of Kimatuumbi phonology have been discussed in various publications, for instance Cowper and Rice (1987), Hayes (1990), Odden (1987, 1990), Truckenbrodt (1999), focusing particularly on theoretical consequences of certain phrasal sandhi rules of the language, as well as Rosenthall (1994) discussing vowel sequences. Odden (1996) provides a comprehensive overview of the morphology and phonology of the language.

Segmental Inventory

The 'core' segmental phonemes of the language are as follows:

(1)	p	t	č	k	i		u
	b			g	ɪ		ʊ
	m	n	ñ	ŋ	ɛ		ɔ
	w	l	y			a	

Additional phonemes with restricted distribution exist. Surface [d] is common but nearly always preceded by a nasal, e.g., [kyʊ́undɔ̀] 'knot' where it is in complementary distribution with [l]. Since *d* exists in a few words not preceded by a nasal, e.g., [daámu] 'blood,' [diíwai] 'wine,' it must be granted marginal phonemic status. The voiced consonants written *b, d, g* are pronounced as implosive [ɓ ɗ ɠ] when not preceded by a nasal, thus more narrowly [ɗaámu], [ɓɔ̀na] 'to see,' [ɠálaambuka] 'to change.' Phonetic [ǰ] exists only after a nasal ([ñǰíiŋgiilɛ] 'I entered'), in complementary distribution and morphophonological alternation with the implosive palatal [ʄ] ([ʄ'iiŋgya] 'to enter.') Surface [ʄ'] itself is almost always limited to root-initial position before the vowel [i], and there it is in free variation with [y], hence [yíiŋgya] 'to enter.' It does also appear elsewhere in a few loan words, such as [ʄ'aáʄ'i] 'judge.' Other restricted phonemes are [v], [f], [z], and [r], which only appear in loanwords ([zaábu] 'gold,' [kɔlɔfiíndɔ] 'walking stick,' [ɗɛrɛ́va] 'driver.'

The consonant transcribed as [č] may be pronounced [č], [tˢ] or [s]. Some speakers appear to always use [s], whereas others tend more toward [č]. In addition, there are lexical preferences (but not requirements) for [s] vs. [č], where [suuŋgulá] 'rabbit' tends to be pronounced with [s] and [čaáŋgu] 'mine' tends to have [č]: but, [čuuŋgulá] and [saáŋgu] are sometimes used. Words such as [suuŋgulá] are mostly loanwords and reflect the original Swahili consonant (*sungura*).

To simplify transcriptions, the implosive/nonimplosive variants are not noted because they are predictable, and the voiced palatal obstruent is written [j]. Since there is no contrast between [e o] and [ɛ ɔ], the mid vowels are written as [e o]. For the sake of comparing data presented here with that published elsewhere, [ɪ ʊ] here correspond to [i u] elsewhere, while [i u] here correspond to [i̧ u̧] used elsewhere, the latter using the traditional Bantuist diacritic for the 'superhigh' vowels. Finally [ŋ] will be used here rather than [n'] found elsewhere.

Vowel length is distinctive, and is indicated by doubling the vowel. The language has H and L tones, which phonologically functions as a H-toned versus toneless distinction. The interpretation of tone marks is as follows, using [a] to represent any vowel:

(2)	á	a	aa	áa	aá
	short H	short L	long L	long fall	long rise

All contour tones appear on long vowels, and there are no level-H long vowels except in one lexical item [sáána] 'very,' which derives from Swahili, and certain cases of vowel contractions.

Syllable Structure

Syllables in Kimatuumbi never have a coda consonant, except in a few loan words ([boóksi] 'box,' [bakteéliya] 'bacteria'). There are also no diphthongs, i.e., bivocalic syllable nuclei, in the language: vowel sequences such as [ui], [oi], [ai], [ao], etc., are disyllabic, thus [tá.u.na] 'to chew' has three light syllables. There are numerous lines of evidence showing the lack of diphthongs discussed in Odden (1996). One relevant example is that such sequences do not trigger the tone rule Heavy Retraction where /u.kaa.té/ becomes [u.kaá.te] 'you should cut,' as shown by [u.ta.u.né] 'you should chew,' not *[u.ta.ú.ne], since there is no preceding heavy syllable.

Consonant onsets conform to the pattern NCG, N standing for nasal, C standing for any consonant, and G standing for glide. Any of these elements is optional. The following data illustrate the onset possibilities.

(3)	No onset	áka	'net-hunt'
	C-onset	bendeéla	'flag'
		yʊ́ba	'hide'
		maána	'meaning'
	CC-onset	bwiími	'greed'
		kyaátí	'valley farm'
		mbaángo	'warthog'
		mméya	'white ant'
		mwaábí	'sorcerer'
	CCC-onset	mbweégá	'side'

Segmental Phonology

The main segmental processes of Kimatuumbi pertain to N + C sequences and vowel height harmony: additional processes of *l*-insertion, *b*- and *g*-deletion, and glide deletion are discussed in Odden (1996, Sect. 3.3–7).

NC Interactions

Kimatuumbi is somewhat unusual among Bantu languages in having two distinct sets of phonological alternations affecting nasal plus consonant sequences, which ultimately relates to the derivational source of the NC sequence. NC sequences arise directly by morpheme concatenation involving the class 9–10 noun prefix /ñ/ or via a rule optionally reducing the 1sg pronoun prefix /ni/ to *n*, triggering changes which are referred to as *n*-type alternations. A partially distinct set of alternations arise from the optional or obligatory reduction of various prefixes of the form /mu/ to just a nasal, which are referred to as *mu*-type alternations. The different processes are summarized in (4).

(4)

underlying *consonant*	n	mu — *Interaction type*
p,t,č,k	[mb,nd,ñǰ,ŋg]	[mp,nt,ñč,ŋk]
b,d,l,j,g,l	[mb,nd,nd,ñǰ,ŋg]	[mm,nn,nn,ññ,ŋŋ]
m,n,ñ,ŋ	[m,n,ñ,ŋ]	[mm,nn,nn,ññ,ŋŋ]
w,y	[ŋgw,ñǰ]	[ŋŋw, ññ]

Common properties transcending the two nasal classes include place assimilation, and approximant-hardening where [w l y] become [gw, d, ǰ]. The types of nasals are differentiated in terms of whether they trigger postnasal voicing (the *n*-type nasal does), postnasal nasalization (the *mu*-type nasal does), or degemination (the *n*-type nasal does). Examples of these processes are given below: n-type alternations are illustrated using the 1sg perfective verb and the mu-type use the 2pl perfective: the optional unreduced 1sg and 2pl variants motivate the underlying consonant.

(5)

n-type *1sg subject*		mu-type *2pl subject*		
ni-téliike	n-déliike	mu-téliike	n-téliike	‘cook’
ni-bwéeni	m-bwéeni	mu-bwéeni	m-mwéeni	‘see’
ni-nóolite	nóolite	mu-nóolite	n-nóolite	‘sharpen’
ni-wíile	ŋ-gwíile	mu-wíile	ŋ-ŋwíile	‘die’

Vowel Harmony

As in most Bantu languages, vowel coocurrence is most free in root-initial position, and is highly restricted within the stem thereafter. Only underlying /a i u/ are found in postinitial position, and are subject to height harmony whereby /i/ becomes [e] after [o e], [ɪ] after [ɪ ᴜ], and /u/ becomes [o] [o] (but not [e], and [ᴜ] appear after [ɪ ᴜ]. The high vowels [i u] then appear elsewhere. An illustration of the pattern with the causative suffix /iy/ is seen below.

(6)

plain	*causative*	
čáag-a	čáag-iy-a	‘grind’
ín-a	in-iy-a	‘dance’
kᴜ́n-a	kᴜ́n-iy-a	‘grate’
číinj-a	číinj-iy-a	‘slaughter’
ᴜ́ᴜg-a	ᴜ́ᴜg-Iy-a	‘bathe’
bóol-a	bóol-ey-a	‘de-bark’
čéeŋg-a	čéeŋg-ey-a	‘build’

Examples with the vowel /u/ harmonizing in the second syllable include [bᴜ́tuka] ‘run,’ [típuka], ‘pound to flour’ [čápuka] ‘get soiled,’ [tᴜ́ᴜmbᴜka] ‘fall,’ [pílᴜka] ‘trip over,’ [épuka] ‘avoid’ [bómoka] ‘be destroyed.’

Hiatus Resolution

Like many Bantu languages, Kimatuumbi has phonological processes that eliminate vowel sequences, but unlike a language such as Kikerewe, which has absolutely no vowel sequences, elimination of vowel hiatus in Kimatuumbi is only partial. Many of the hiatus-eliminating strategies in the language are optional or are subject to restrictions on morphosyntactic domain: ultimately, quite a number of V + V sequences remain. One of the interesting concomitants of these rules is that the remaining vowel is compensatorily lengthened.

Glide formation affects prevocalic high vowels, and can be easily motivated in verb paradigms by combining subject or object prefixes plus a vowel-initial vs. consonant-initial roots.

(7)

Infinitive	*3p-V-past*	*1s-V-past*	*1p-V-past*
čón-a	ba-čón-ite	ni-čón-ite	tu-čón-ite
éleew-a	ba-éleew-ite	ny-eéleew-ite	tw-eéleew-ite
ᴜ́bᴜk-a	ba-ᴜ́bwiik-e	ny-ᴜᴜ́bwiik-e	tw-ᴜᴜ́bwiik-e
ᴜ́ᴜg-a	ba-ᴜ́ᴜg-ite	ny-ᴜ́ᴜg-ite	tw-ᴜ́ᴜg-ite
ák-a	ba-ák-ite	ny-aák-ite	tw-aák-ite
áam-a	ba-áam-ite	ny-áam-ite	tw-áam-ite
	2s-V-past		
	u-čón-ite	‘draw’	
	w-eéleew-ite	‘understand’	
	w-ᴜᴜ́bwiik-e	‘climb’	
	w-ᴜ́ᴜg-ite	‘bathe’	
	w-aák-ite	‘net-hunt’	
	w-áam-ite	‘emigrate’	

Glide formation is subject to a number of conditions. For example, it is optional in domain-medial syllables (see ‘Organization of Rules’ below), it does not affect long vowels (Odden, 1996: Sect. 4.1.1), it has a few lexical exceptions, and has complex tonal conditioning where H toned vowels do not undergo glide formation before a long vowel.

The second main hiatus-reduction process is Vowel Contraction, where /a+i/ and /a+u/ optionally become [ee] and [oo]. This rule only applies between prefixes or proclitics and prefixes, never involving a root-initial vowel. Examples of this involving /a/ in a subject or tense-aspect prefix, plus /i/, /u/ in a subject or object prefix can be used to motivate this process.

(8)

uncontracted	*contracted*		
ba-i-káatite	beekáatite	(ba- ‘they’, -i-‘it’)	‘they cut it (cl. 9)’
ba-u-káat-ite	bookáatite	(-u- ‘you’)	‘they cut you’
ka-i-bolá	keebolá	(ka- ‘when’)	‘when it (cl. 9) rotted’
ka-u-bolá	koobolá	(-u- ‘it (cl. 3)’)	‘when it (cl. 3) rotted’

Root-initial vowels do not undergo contraction, so [a-ínite] ‘he danced’ never becomes *[eénite]. Long [aa] also cannot contract: cf. [naa-i-káatite] ‘I cut it,’ *[nee-káatite].

Also falling in the category of hiatus-resolving strategies is the desyllabification of an underlying or derived syllabic nasal, with compensatory lengthening of the preceding vowel. Although most Bantu languages with contrastive vowel length have predictably long vowels before NC clusters, Kimatuumbi no

longer has such lengthening as part of its active phonology. Within roots, vowels are generally long before NC sequences, for example *áandika* 'to write,' *kύυmbwa* 'to beat,' *lisaambalá* 'hide.' Because vowel length within the stem is restricted to (approximately) the last three moras, short vowels are found before NC in words such as *yánjaaya* 'to turn something drying,' *simbasiímba* 'road,' *yύmbItIka* 'to be quick,' *lipúndugulu* 'fog.' Even so, there are cases of short vowels before NC within the last three moras: *kičénjele* 'cock's comb,' *lukóngobe* 'hardwood,' *nčongóma* 'thorn-tree,' *ntíndí* 'buttermilk.' There is also no active lengthening across morpheme boundaries, thus the noun class prefix vowel is not lengthened in *ki-ndoólo* 'sweet potato,' *ka-ngápa* 'little bog,' *tu-ndó* 'little buckets,' nor is there any lengthening at the phrasal level (*mbusi ngύlύ 'big grating stool,' kila ndó 'every bucket')*.

However, there is the **appearance** of lengthening before a class of NC clusters, namely those involving reduction of /mu/. As argued in Odden (1996: Sect. 4.1.7–8), this is because as an intermediate step in the derivation, /mu/ reduces to a syllabic (moraic) nasal, which then transfers its mora to a preceding vowel if there is one. The vowel-lengthening concomitant of reduction of *mu-* can be seen in (9). The examples in (9a) involve phrasal combinations of a verb plus a locative where the locative prefix /mu/ optionally reduces, and the examples in (9b) involve lengthening of the prefix *ka-* 'when' before the 2pl subject /mu/.

(9a)	agonjite mύ-kilíndıılo	~ agonjiteé ŋ-kilíndıılo
	'he slept in the guard house'	
	naatíkwiibiikámumipύkó	~naatíkwiibiikáammipúkó
	'I put them in the bags'	
(9b)	kamuteleká	~ kaanteleká 'while2pl.werecooking'
	kamukalaáŋga	~kaaŋkalaáŋga 'while 2pl. were frying'

Tone

Bantu languages typically have a number of restrictions on the surface and underlying placement of tone in a word, and Kimatuumbi imposes more than the usual distributional restrictions on words. A difference must be made between noun tone, which is lexical, and verb tone, which is purely predictable by reference to phonological and morphological factors.

Nouns

With very few exceptions, noun stems have at most a single underlyingly H-toned syllable, and the underlying position of H is noncontrastive within bimoraic syllables. What is striking about noun tone in Kimatuumbi is that on the surface, this highly restricted nature of tonal specifications is not obvious. Typical examples illustrating the divergence between underlying and surface tones are given in (10).

(10)	mpaká	'boundary'	/mu-paká/
	ntíndí	'buttermilk'	/mu-tíndi/
	mikóta	'canes'	/mi-kota/
	mbaká	'cat'	/ñ-paká/
	ndími	'tongues'	/ñ-límí/
	ñáma	'meat'	/ñama/
	ŋkaáte	'loaf'	/mu-kaaté/
	ntóopé	'handle'	/mu-tóope/
	mpúuŋga	'paddy'	/mu-puuŋga/
	mboópo	'machete'	/ñ-poopó/
	ndooñá	'lip-rings'	/ñ-tóoña/
	ŋguúbe	'pig'	/ñ-kuube/
	masípitaáli	'hospitals'	/ma-sipitaalí/
	nčímbılíılí	'drip'	/mu-čımbılíılı/
	sipitaáli	'hospital'	/sipitaalí/
	njeénjemá	'mosquito'	/ñ-čeenjéma/
	matóondolo	'steinbucks'	/ma-tóondolo/
	ndoóndolo	'steinbuck'	/ñ-tóondolo/
	kipıíkııtí	'piece of wood'	/ki-pııkíıtı/

One of the sources of opacity in the noun tone system is that while all nouns have a H tone in their citation form, the same is not true at the phrasal level. Compare the examples in (11a) with those in (11b).

(11a)	*N*	*'it's not N'*	
	ki-wíkılyo	ŋga ki-wıkılyo lí	'cover'
	mi-kóta	ŋga mi-kota lí	'sugar canes'
	ki-kóloombe	ŋga ki-koloombelí	'shell'
	ŋ-guύŋguni	ŋga ŋ-guuŋguni lí	'bedbug'
(11b)	lu-kóŋgobe	ŋga lu-kóŋgobe lí	'wood'
	ma-tógolo	ŋga ma-tógolo lí	'waterbucks'
	ki-kálaaŋgo	ŋga ki-kálaaŋgo lí	'pan'
	diíwai	ŋga diíwai lí	'wine'

Nouns such as the first set have no underlying H tone and those in the second group have an underlying H. This distinction is neutralized prepausally, by a rule assigning H to the second mora of the word to any toneless word in prepausal position. Words such as [ki-kóloombe], which have a CV noun class prefix, have their H on the first stem syllable: nouns with no moraic prefix such as [ŋ-guúŋguni] are what show that the correct generalization is 'second vowel of the word.'

Two rules assign additional H tones to a noun beyond the underlying lexical tone, whose operation contradicts the surface generalization that stems have only a single H. First, H is assigned to the root-initial mora when a mora-bearing noun class prefix – other than the class 5 prefix /li-/ – precedes the stem and the lexical H is not on the following syllable. A number of nouns allow both nonmoraic and moraic prefixes, which results in alternations in the tone of the root-initial syllable.

(12)	sipitaáli	‘hospital’	ma-sípitaáli	‘hospitals’
	bisikúti	‘biscuit’	ka-bísikúti	‘little biscuit’
	čwaakalá	‘lime’	ma-čwáakalá	‘pieces of lime’
	solokó	‘black beans’	lu-sólokó	‘black bean’

Examples such as *ma-paká* ‘cats,’ *ma-poópo* ‘machetes,’ where the root-initial mora precedes a H tone, illustrate the necessity of a toneless vowel between the target of H-insertion and the lexical H in the stem.

The second source of multiple H tones inside a stem is a Tone Copy rule that assigns H tone to the final vowel after a penultimate H tone. This rule only applies prepausally, so in certain phrase-medial positions, the final H will be lacking (because of a phrasal tone rule assigning H, discussed in ‘Phrasal Tone Insertion’ below, the typical phrasal context where the final H is surface absent is VP-medial).

(13)	lukópé	lukópe ...	‘eyelid’
	lulímí	lulímɪ ...	‘tongue’
	kikwɪkwíɪkwɪ	kikwɪkwíɪkwɪ ...	‘belch’
	lukóonjí	lukóonji ...	‘finger’
	lutebéelé	lutebéelé ...	‘amaranthus’

Besides restricting underlying H tones to one-per-stem, there is also no underlying contrast between rising and falling tones, even though such a distinction is made on the surface. In nouns, most rising tones come from the rule Heavy Retraction, where a H tone shifts leftward to a long vowel forming a rising tone. Since a phrasal sandhi rule shortens long vowels in phrase-medial position (see ‘Shortening’ below), the underlying position of the (unretracted) H tone can be detected relatively easily by including a modifier that induces vowel shortening.

(14)	mabiíbi	/ma-biibí/	‘grandmothers’
	kiboóko	/ki-bookó/	‘stick’
	asíkaáli	/a-sikaalí/	‘soldiers’
	kipílɪɪ́lɪ	/kipílɪɪlɪ́/	‘scorpion’
		mabibí gaángu	‘my grandmothers’
		kibokó kilaáso	‘long stick’
		asíkalí awaáá	‘dead soldiers’
		kipílɪlɪ́ɪ ŋkʊ́lʊ́	‘big scorpion’

In the examples above, the retracted H is underlyingly on the word-final vowel. The H may also be on a penult that is preceded by a long vowel, as in (15). In that case, the H is both subject to spreading to the final syllable by Tone Copy, and also is retracted.

(15)	ligaándamá	/ligaandáma/	‘spleen’
	nguúngutá	/nguungúta/	‘bean’
	lutʊʊ́tʊlʊ́	/lutʊʊtʊ́lʊ/	‘popcorn’
	kipɪɪ́kɪɪtɪ́	/kipɪɪkɪɪ́tɪ/	‘wood’
		ligandáma liikʊ́lʊ́	‘big spleen’
		ngungúta yaángu	‘my bean’
		lutʊtʊ́lu lwaángu	‘my popcorn’
		kipɪɪkɪ́ɪtɪ kikʊ́lʊ́	‘large wood’

Verb Tone

The overwhelming majority of Bantu languages have a contrast between H-toned and L-toned or toneless stems, expressed underlyingly by a two-way tonal contrast on the first root syllable. Kimatuumbi is one of a small number of tonal Bantu languages that have lost verbal lexical tone distinctions – Yao, Makonde, Makua, Kikuria, and Olusaamia are other examples – and instead assign surface tone based purely on the verb’s tense-aspect category. In Kimatuumbi verbs, a H is assigned to a particular position – stem-initial, stem-final, second-stem vowel, or third-stem vowel – depending on inflectional category.

The pattern of most basic verb forms is that H is assigned to the first-stem vowel. This pattern is used in the infinitive, perfective, future, and certain verb-focal forms.

(16)	*infinitive*	*3p perfective*	*1pl future*
	lʊ́ba	ba-lʊ́bite	twaa-lʊ́ba
	báamba	ba-báambite	twaa-báamba
	čáŋgaala	ba-čáŋgaalite	twaa-čáŋgaala
	pápaakiya	ba-pápaakiiye	twaa-pápaakiya
	pákatika	ba-pákatiike	twaa-pákatika
	láambata	ba-láambiite	twaa-láambata
	káčabanika	ba-káčabaniike	twaa-káčabanika
	kéeŋgeemba	ba-kéeŋgeembite	twaa-kéeŋgeemba
		3s strong verb-focus	*prog*
		eendánoo-lʊ́ba	‘ask’
		eendánoo-báamba	‘top’
		eendánoo-čáŋgaala	‘wonder’
		eendánoo-pápaakiya	‘grope’
		eendánoo-pákatika	‘shake down’
		eendánoo-láambata	‘lick’
		eendánoo-káčabanika	‘be confused’
		eendánoo-kéeŋgeemba	‘dig up’

Strong verb-focus forms such as [eendánoolʊ́ba] are best translated as ‘he is even *asking*.’ By comparison, the weak verb focus forms, which lack the prefix *-noo-*, assign H tone to the final vowel of the stem.

(17)	*3s pres. prog. verb-focus*	*3p past verb-focus*	
	eendaá-lʊbá	baatí-lʊbá	‘ask’
	eendaá-baambá	baatí-baambá	‘top’
	eendaá-čaŋgaalá	baatí-čaŋgaalá	‘wonder’
	eendaá-papaakiyá	baatí-papaakiyá	‘grope’
	eendaá-pakatiká	baatí-pakatiká	‘shake down’
	eendaá-laambatá	baatí-laambatá	‘lick’
	eendaá-kačabaniká	baatí-kačabaniká	‘be confused’
	eendaá-keeŋgeembá	baatí-keeŋgeembá	‘dig up’

In most dependent clause verb tenses (relative clauses, when-clauses, and ‘without V-ing’ clauses), a H tone is assigned to the second stem vowel.

(18)	*when we V*	*without V’ing*	
	patʊ́-lʊbá	wanga-lʊbá	‘ask’
	patʊ́-baámba	wanga-baámba	‘top’
	patʊ́-čaŋgáala	wanga-čaŋgáala	‘wonder’

patʊ́-papáakiya	wanga-papáakiya	'grope'
patʊ́-pakátika	wanga-pákatika	'shake down'
patʊ́-laámbata	wanga-laámbata	'lick'
patʊ́-kačábanika	wanga-kačábanika	'be confused'
patʊ́-keéŋgeemba	wanga-keéŋgeemba	'dig up'

Past tense subordinate verbs as well as deverbal stative adjectives have an additional H tone on the stem-final vowel.

(19)	'when you V'd'	'in a V'd state (cl. 7)'	
	paʊ́-lʊbíté	ki-lʊbá	'ask'
	paʊ́-baámbité	ki-baámbá	'top'
	paʊ́-čaŋgáalité	ki-čaŋgaálá	'wonder'
	paʊ́-papáakiiyé	ki-papáakiyá	'grope'
	paʊ́-pakátiiké	ki-pakátiká	'shake down'
	paʊ́-laámbiité	ki-laámbatá	'lick'
	paʊ́-kačábaniiké	ki-kačábaniká	'be confused'
	paʊ́-keéŋgeembite	ki-keéŋgeembá	'dig up'

H tone is assigned to the first-stem vowel in the subjunctive and participial tenses.

(20)	'you should V'	'while I was V-ing'	
	u-lʊbé	kani-lʊbá	'ask'
	u-baámbe	kani-baámba	'top'
	u-čaŋgaále	kani-čaŋgaála	'wonder'
	u-papaakíye	kani-papaakíya	'grope'
	u-pakatíke	kani-pakatíka	'shake down'
	u-laambáte	kani-laambáta	'lick'
	u-kačabánike	kani-kačabánika	'be confused'
	u-keeŋgéembe	kani-keeŋgéemba	'dig up'

Forms such as [u-baámbe] and [kani-baámba] derive from intermediate *ubaambé* and *kanibaambá*, H tone is assigned to the third *stem* vowel, but the H shifts left by Heavy Retraction. One difference between nouns and verbs in the application of heavy Retraction is that the rule will shift a H to a prepenult long syllable in a noun ([kipɪ́kɪɪtí] ← /kipɪɪkɪ́ɪtɪ/ 'wood') but not in a verb (u-papaakíye 'you should grope').

The final tone pattern used in the verb system assigns H tone to the first and third vowels of the stem, in deverbal nominalizations selecting the suffix *-i*.

(21)	n-nʊ́bí	'asker'	m-máambí	'one whotops'
	ŋ-kógotí	'scraper'	n-čáŋgaáli	'wonderer'
	n-nóŋgeláaŋgi	'converser'	m-pápaákiyi	'groper'
	m-pákatíki	'shaker'	n-náambáti	'licker'
	ŋ-káčabániki	'confused person'	ŋ-kéeŋgéembi	'digger'

Phrasal Phonology

Kimatuumbi has a rich set of phonological alternations applying at the phrasal level. Some of these are utterance level rules which only care whether a word is at the end of an utterance or not, while others are sensitive to the specific syntactic relations between the involved words.

Shortening

One sandhi rule is Shortening, which shortens long vowels within a stem when one word is followed by another word, provides that they are in a head-modifier relationship. Examples of this involving nouns and adjectives are seen in (22), and verbs followed by objects or adverbs are seen in (23).

(22)	ki-kóloombe	'cleaning shell'
	ki-kólombe čaáŋgu	'my cleaning shell'
	ki-kólombe kikʊ́lʊ́	'big cleaning shell'
	li-peésa	'money'
	li-pesá lyé	'the money'
	li-pesá lyanáalyoobíté	'money that I lost'
(23)	naa-kálaaŋgite	'I fried'
	naa-kálaŋgite mapíiŋgá	'I fried eggs'
	naa-kálangite líiso	'I fried yesterday'
	u-paánde	'you should plant'
	u-pandé maloómbe	'you should plant corn'
	u-pandé paapasá	'you should plant the day after tomorrow'
	u-pandé lí	'you should not plant'

Shortening interacts with verb focus, so that a focused verb does not undergo shortening.

(24)	neendá-twɪɪká misígó	'I'm *lifting* the loads'
	aatí-pwaatá Mambóondo	'he *followed* Mamboondo'
	tweendeé-kuŋuundá lí	'we aren't *filtering*'
	toola lusólokó	'just *take* the bean!'

If the following word is not a modifier of the previous word, there is no shortening. For example, inside the VP, one noun does not undergo shortening when followed by another word outside of the noun's phrase; nor does an adjective undergo shortening before another adjective.

(25)	naa-m-péi	Mamb**oo**ndo	eéla	'I gave M. money'
	I-him-gave	*Mamboondo*	*money*	
	naa-ki-bwéni	kikól**oo**mbe	líilí	'I didn't see a shell'
	I-it-saw	*shell*	NEG	
	mbopó	nd**aá**so	ngulʊ́	'big long machete'
	machete	*long*	*big*	

Verbs do not undergo shortening when followed by a sentential adverb, which is outside of the verb's phrase; nor is there shortening in the verb of a subordinate clause that is followed by a word of the matrix clause.

(26)	aa-kálaaŋga	sáána		'he will really fry'
	he.fut-fry	*really*		
	naa-m-péi	ywaá-kaátite	eéla	'I will give money to the one who cut'
	I-him-gave	*who.he-cut*	*money*	

Phrasal Tone Insertion

Another syntactically conditioned phonological rule of Kimatuumbi is Phrasal Tone Insertion, which adds a H tone between maximal projections that are dominated by a maximal projection:

(27) $[_{Xmax} \ldots [_{Ymax} __] [_{Zmax}] \ldots]$

This configuration rule is found in a number of contexts, for example between a subject NP and a VP.

(28) Mambóondo (/Mamboomdo/) P.N.
Mamboond**ó** aawíile 'Mamboondo died'
kiyógoyo 'bird (sp.)'
kiyógoy**ó** čaatí tʊʊmbʊká 'the bird fell'

H tone is inserted at the end of each phrase coming before the verb, which may include objects and adverbs owing to the free word order of the language.

(29) iíjuma 'Friday'
iíjum**á** kiyógoy**ó** Mamboond**ó** naampéi lí
'I didn't give Mamboondo a kiyogoyo on Friday'

Tone insertion only applies to the end of the phrase that is a sister to the VP; thus an adjective inside the subject NP does not receive any H.

(30a) mund**uu** ntokómaʊ́ a-wíile 'a sluggish person is dead'
person sluggish he-be dead
*mund**úu** ntokóma**ú** awíile

NP and sentential conjunctions are also subject to this rule.

(30b) patúkatítee ñam**á** Ukumʊ́ akuŋundite wɪ́ímbɪ
when-we-cut mean ʊ. he-filtered beer
'when we cut the meat, ʊkumu filtered the beer'
mpuuŋg**á** na kindoólo 'rice and sweet potato'
rice and sweet potato
kikálaaŋgó aʊ́ kibIgá 'frying pan or a pot'
frying pan or pot

The juncture between a VP and a sentential adverb provides another context for this rule.

(31) niimpéendiile 'I like him'
niimpéendiil**é** píta 'I really like him'
aakálaanga 'he will fry'
aakálaang**á** kwaáli 'perhaps he will fry'

Initial Tone Insertion

Two of the syntactically conditioned rules are sensitive to crossing sentence boundaries. One of these rules, Initial Tone Insertion, places a H at the beginning of certain words – a relatively arbitrary collection of adverbs, function words, and agreement prefixes, plus nouns beginning with a long vowel – when the preceding stem has no medial H.

(32) kindoló **ča**-Mambóondo 'sweet potato of Mamboondo'
kiwɪkɪlyo **čá**-Mambóondo 'cover of Mamblondo'
tu-tyatyakíkiye **ma**laáʊ 'we should make plaster tomorrow'
twaa-tyátyakikiya **má**laáʊ 'we will make plaster tomorrow'
paní-teléka **lii**yawá 'when I cook pumpkin'
ni-teleka **líi**yawá 'I cook pumpkin'

Unlike Shortening, the interacting words do not need to be in the same phrase, thus H can be assigned to the instrumental prefix *na-* when preceded by a toneless direct object. H tone can also be assigned to a word in a subordinate clause, conditioned by the preceding word in the matrix clause.

(33) naammʊ́lige Kiwiiyo **ná**-mboópo 'I killed Kiwiyo with a machete'
I-him-killed Kiwiiyo with-machete
naammʊ́lige **ká**apangá kaási 'I killed him when he was working'
I-him-killed when-he-do work

But no H is inserted when the trigger word is in a subordinate clause and the target is in the higher clause.

(34) aambákiye pánga twaabʊ́ʊka kindaái
he-me-told that we-will-return tomorrow
'He told me that we will return tomorrow'
*aambákiye pánga twaabʊ́ʊka kíndaái

As noted in Odden (1996: Sect. 6.1.4), this motivates the condition that the trigger command the target, in the traditional sense of command.

Lengthening

A second sentence-level rule with a clause-mate condition is Lengthening, which lengthens a vowel before a CV(C)V noun stem. Examples of this are seen in (35).

(35) aatweti ñoóme 'he took a cat'
aatwéti ndó 'he took a bucket'
aatwétii ñáma 'he took a bottle'
aatwétii píli 'he took a puff adder'

This rule also applies to the class 5 noun prefix /li-/, but not any other noun class prefixes.

(36) lii-yawá 'pumpkin' ma-yawá 'pumpkins'
lii-kʊ́tI 'coconut leaf' ka-kʊ́tí 'little coconut leaf'

Failure of Lengthening to cross a sentence boundary is seen below:

(37) ni-yuwine $[_s$ mbaká awíile 'I heard that the cat died'
I-heard cat it-died

The relevance of internal sentence boundaries is not a widely discussed aspect of the syntax–phonology interface problem, and it is not clear how prosodic theories of the interface would handle the relevance of the node S in phonology. See Odden (1996: Sect. 6.2) for discussion involving the notion 'pause.'

Syntactically Unconditioned Rules

There are, in addition, a number of other external sandhi rules that apply without consideration of syntactic factors, either applying between any two words or applying just at the end of the utterance. One example of that is the rule H Tone Assignment, which assigns a H tone to the second vowel of any toneless stem, when it stands at the end of the utterance; this rule was discussed in 'Nouns' above. Similarly, Tone Copy, which spreads a penult H tone to the final vowel of the word in nouns, only applies to prepausal nouns. A third rule is Perfective Truncation, a rule that optionally deletes the final syllable *le* in a perfective verb which is not prepausal.

(38) aantélekile Mambóondo 'he cooked for Mamboondo'
aantéleki Mambóondo 'he cooked for Mamboondo'
baakʊ́ñwile Libulʊ́le 'they touched Libulule'
baakʊ́ñwi Libulʊ́le 'they touched Libulule'

Other such rules are discussed in Odden (1996: Sect. 6.2).

Organization of Rules

Interaction between phonological rules in Kimatuumbi can be rather complex and challenges certain theoretical assumptions regarding lexical vs. postlexical rules made in the theory of Lexical Phonology. The interaction between Glide Formation, morphosyntactically governed rules of tone-assignment in verbs, and the sandhi rule Shortening especially result in a considerable degree of opacity.

The rule Glide Formation applies cyclically, where a cyclic domain break is defined at the junction between proclitics vs. the basic word. This can be demonstrated by comparing analogous vowel sequences, differentiated as to whether the first is in a proclitic or not. One class of examples involves /CV-CV-V/ sequences morphologically distinguished as locative prefix + noun class prefix + vowel initial root, versus subject prefix + object prefix + vowel initial root. One condition on Glide Formation is that it applies obligatorily to a prevocalic high vowel that is initial – in the domain visible at that stage in the cycle – but is optional when applied to a domain-medial prevocalic high vowel.

(39) /a-lu-áandiike/ → [aluáandiike] ~ [alwáandiike] 'he wrote it'
he-it-wrote
/pa- [lu -áanjʊ́]/ → [palwáanjʊ́] 'in the firewood'
at -n. class -firewood

In the case of /a-lu-áandiike/, both the subject prefix *a-* and the object prefix *-lu-* are underlyingly present at the stage where Glide Formation applies, and since *lu* is not initial in that domain, Glide Formation is optional. In contrast, in the case of /pa-[lu-áanjʊ́], first the word [lwáanjʊ́] is derived from /lu-áanjʊ́/, and at that stage *lu* is initial in the cyclic domain, this Glide Formation is obligatory. Subsequently, the proclitic /pa/ is added, but this has no effect on the optionality of Glide Formation on the earlier cycle.

The sandhi rule Shortening applies both before Glide Formation (applied at the stem level) and after Glide Formation (at the word level). The consequence of this interleaving of Shortening between applications of Glide Formation is that long vowels that are created via the compensatory lengthening side effect of Glide Formation at the stem level (*akyaana*) are subject to Shortening when a modifier follows, but long vowels created via Glide Formation on the poststem cycle (*twaakyaana*) are not subject to shortening. Phrasal Shortening of long vowels created at the stem level is seen in (40).

(40) /ak-i-ana/ → [ákyaana] 'to net-hunt for each other'
hunt-for-e.o
/ak-i-ana itʊ́ʊmbIlI/ → [akyana itʊ́ʊmbIlI] 'to net-hunt monkeys for each other'
hunt-for-e.o monkeys

In (41), nonshortening of long vowels created at the word level is evident.

(41) /tu-ak-i-ana itʊ́ʊmbIlI/ → [twaakyana itʊ́ʊmbIlI]
we-hunt-for-e.o monkeys
'we net-hunt monkeys for each other'

This poses a problem for Lexical Phonology, in that the presumably postlexical rule Shortening is sandwiched between applications of the cyclic rule Glide Formations.

The rule Heavy Retraction shifts a H tone from a final syllable to a long penult, so that predicted/ u-toolé/ 'you should take' – cf. [u-teleké] – becomes [u-toóle]. But a long vowel derived via Glide Formation does not trigger tone retraction.

(42) áka 'to hunt' w-aaké 'you should hunt'
éka 'to laugh' w-eeké 'you should laugh'

Thus Glide Formation counterfeeds Heavy Retraction.

Since root-initial vowels can be underlyingly long, this results in particularly opaque alternations in (43), where Glide Formation is not necessary to bring about the long vowel triggering Retraction. In these examples, Retraction applies to the H that is expected to be on the final syllable.

(43)	áama	'to emigrate'	w-aáme	'you should emigrate'
	íɪmba	'to dig'	w-íɪmbe	'you should dig'

Thus, underlying /u-aamé/ undergoes Retraction to give *uaáme*, then Glide Formation resulting in surface *waáme*. The decision whether H tone retracts thus is based on the distinction between long vowels derived as a side effect of Glide Formation, vs. those that are underlyingly long vowels.

Opaque application of Shortening is seen in (44) where the verb is followed by an object.

(44)	/u-aamé lí/	→ [waamé lí]	'you shouldn't emigrate'
	/u-Πmbé makéne/	→ [wΠmbé makéne]	'you should dig holes'

The evidence that Shortening has actually applied in [waamé lí] is indirect, since there is no surface length alternation. The variation between [waáme] and [waamé ...] is parallel in terms of the underlyingly similar forms [ukaáte] 'you should cut' and [ukaté ...]. Thus, to explain the alternating location of H, we presume that /u-aame lí/ first undergoes Shortening and tone assignment, giving *u-amé lí* (analogous to *ukaté lí* 'you should not cut') and since in this form there is no long vowel, Heavy Retraction does not apply. Afterwards, Glide Formation applies to give [waamé lí], effectively restoring the original vowel length, this such an A → B → C derivation is a classic example of the 'Duke of York Gambit' (Pullum, 1976).

See also: Bantu Languages; Lexical Phonology and Morphology; Tone in Connected Discourse; Tone: Neurophonetics; Tone: Phonology.

Bibliography

Cowper E & Rice K (1987). 'Are phonosyntactic rules necessary?' *Phonology 4,* 185–194.

Guthrie M (1971). *Comparative Bantu* (4 vols). Farnborough: Gregg.

Hayes B (1990). 'Precompiled phrasal phonology.' In Inkelas S & Zec D (eds.) *Proceedings of the Phonology – Syntax Connection Conference*. Chicago: University of Chicago Press. 85–108.

Nurse D (1988). 'The diachronic background to the language communities of SW Tanzania.' *Sprache und Geschichte in Afrika 9,* 15–116.

Odden D (1987). 'Kimatuumbi phrasal phonology.' *Phonology 4,* 13–36.

Odden D (1990). 'Syntax, lexical rules and postlexical rules in Kimatuumbi.' In *Proceedings of the Phonology – Syntax Connection Conference*. Chicago: U. Chicago Press. 59–277.

Odden D (1996). *The phonology and morphology of Kimatuumbi*. Oxford: Clarendon Press.

Odden D (2003). 'Rufiji-Ruvuma (N10, P10–20).' In Nurse D & Philippson G (eds.) *The Bantu languages*. London & New York: Routledge. 529–545.

Pullum G (1976). 'The Duke of York gambit.' *Journal of Linguistics 12,* 83–102.

Rosenthall S (1994). 'Vowel/glide alternations in a theory of constraint interaction.' Ph.D. diss., University of Massachussetts.

Truckenbrodt H (1999). 'On the relation between phonological phrases and syntactic phrases.' *Linguistic Inquiry 30,* 219–255.

Kindaichi, Kyosuke (1882–1971)

A Pérez Pereiro, Georgetown University, Washington, DC, USA

Kyosuke Kindaichi was born on May 5, 1882, in the city of Morioka, Japan. He achieved recognition in Japan and abroad through the study of the Ainu language and, in particular, the yukar, the oral epics of the Ainu, of which he collected many from informants at a time when the Ainu language was already moribund. He is widely considered the founder of modern Ainu studies in Japan.

Kindaichi entered the linguistics department of Tokyo Imperial University in 1904, graduating in 1907 with the dissertation *Particles in the languages of the world*. During this time, he conducted research on the Sakhalin (Karafuto) Ainu and their dialects. In 1913, Kindaichi began teaching Ainu at Tokyo Imperial University, becoming a full professor in 1941. He also accepted a post at Kokugakuin University as professor of linguistics and Japanese in 1922, becoming a professor emeritus in 1968. He continued to conduct fieldwork in Hokkaido through a network of informants. Kindaichi was intrigued by the possibility that the Ainu language might be related to the

Thus Glide Formation counterfeeds Heavy Retraction.

Since root-initial vowels can be underlyingly long, this results in particularly opaque alternations in (43), where Glide Formation is not necessary to bring about the long vowel triggering Retraction. In these examples, Retraction applies to the H that is expected to be on the final syllable.

(43)	áama	'to emigrate'	w-aáme	'you should emigrate'
	íɪmba	'to dig'	w-íɪmbe	'you should dig'

Thus, underlying /u-aamé/ undergoes Retraction to give *uaáme*, then Glide Formation resulting in surface *waáme*. The decision whether H tone retracts thus is based on the distinction between long vowels derived as a side effect of Glide Formation, vs. those that are underlyingly long vowels.

Opaque application of Shortening is seen in (44) where the verb is followed by an object.

(44)	/u-aamé lí/	→ [waamé lí]	'you shouldn't emigrate'
	/u-Πmbé makéne/	→ [wΠmbé makéne]	'you should dig holes'

The evidence that Shortening has actually applied in [waamé lí] is indirect, since there is no surface length alternation. The variation between [waáme] and [waamé ...] is parallel in terms of the underlyingly similar forms [ukaáte] 'you should cut' and [ukaté ...]. Thus, to explain the alternating location of H, we presume that /u-aame lí/ first undergoes Shortening and tone assignment, giving *u-amé lí* (analogous to *ukaté lí* 'you should not cut') and since in this form there is no long vowel, Heavy Retraction does not apply. Afterwards, Glide Formation applies to give [waamé lí], effectively restoring the original vowel length, this such an A → B → C derivation is a classic example of the 'Duke of York Gambit' (Pullum, 1976).

See also: Bantu Languages; Lexical Phonology and Morphology; Tone in Connected Discourse; Tone: Neurophonetics; Tone: Phonology.

Bibliography

Cowper E & Rice K (1987). 'Are phonosyntactic rules necessary?' *Phonology 4,* 185–194.

Guthrie M (1971). *Comparative Bantu* (4 vols). Farnborough: Gregg.

Hayes B (1990). 'Precompiled phrasal phonology.' In Inkelas S & Zec D (eds.) *Proceedings of the Phonology – Syntax Connection Conference.* Chicago: University of Chicago Press. 85–108.

Nurse D (1988). 'The diachronic background to the language communities of SW Tanzania.' *Sprache und Geschichte in Afrika 9,* 15–116.

Odden D (1987). 'Kimatuumbi phrasal phonology.' *Phonology 4,* 13–36.

Odden D (1990). 'Syntax, lexical rules and postlexical rules in Kimatuumbi.' In *Proceedings of the Phonology – Syntax Connection Conference.* Chicago: U. Chicago Press. 59–277.

Odden D (1996). *The phonology and morphology of Kimatuumbi.* Oxford: Clarendon Press.

Odden D (2003). 'Rufiji-Ruvuma (N10, P10–20).' In Nurse D & Philippson G (eds.) *The Bantu languages.* London & New York: Routledge. 529–545.

Pullum G (1976). 'The Duke of York gambit.' *Journal of Linguistics 12,* 83–102.

Rosenthall S (1994). 'Vowel/glide alternations in a theory of constraint interaction.' Ph.D. diss., University of Massachussetts.

Truckenbrodt H (1999). 'On the relation between phonological phrases and syntactic phrases.' *Linguistic Inquiry 30,* 219–255.

Kindaichi, Kyosuke (1882–1971)

A Pérez Pereiro, Georgetown University, Washington, DC, USA

Kyosuke Kindaichi was born on May 5, 1882, in the city of Morioka, Japan. He achieved recognition in Japan and abroad through the study of the Ainu language and, in particular, the yukar, the oral epics of the Ainu, of which he collected many from informants at a time when the Ainu language was already moribund. He is widely considered the founder of modern Ainu studies in Japan.

Kindaichi entered the linguistics department of Tokyo Imperial University in 1904, graduating in 1907 with the dissertation *Particles in the languages of the world*. During this time, he conducted research on the Sakhalin (Karafuto) Ainu and their dialects. In 1913, Kindaichi began teaching Ainu at Tokyo Imperial University, becoming a full professor in 1941. He also accepted a post at Kokugakuin University as professor of linguistics and Japanese in 1922, becoming a professor emeritus in 1968. He continued to conduct fieldwork in Hokkaido through a network of informants. Kindaichi was intrigued by the possibility that the Ainu language might be related to the

Native American languages, noting the polysynthetic, incorporating character of the Ainu verb, which is also found in Amerindian languages.

Several of Kindaichi's students went on to become prominent scholars of the Ainu language themselves. These included Shiro Hattori, a recognized dialectologist of both Hokkaido and Sakhalin Ainu, and Itsuhiko Kubodera, a compiler of Ainu oral literature. Kindaichi was also remarkable for encouraging Ainu people themselves to participate in and conduct research into their own language and oral histories. Among the most important of these scholars were Mashiho Chiri, the first Ainu to attend Tokyo Imperial University, with whom Kindaichi co-authored *Ainu goho gaisetsu* [Outline of Ainu grammar], and Shigeru Kayano, an Ainu anthropologist and politician.

Along with several native speakers of the Ainu language, notably Yukie Chiri and Matsu Kannari, Kindaichi compiled and published various yukar and other stories that might have been lost otherwise. He also helped John Batchelor's adopted daughter, Yaeko Batchelor, publish as part of her crusade against the alcoholism that was ravaging the Ainu community. Although Kindaichi believed the traditional Ainu would disappear, he resisted attempts to characterize the Ainu as a dying race, as was common at the time, and he offered instead the possibility of constructive assimilation – using his star pupil, Mashiho Chiri, as an example of Ainu genius. His work had the double effect of helping to empower a generation of Ainu scholars to become active researchers and interpreters of their culture and of providing a source of pride, in the form of the rich intellectual tradition of the yukar, for a people who for generations had been told that they had no culture or civilization to speak of.

Besides his Ainu studies, Kindaichi was active in the Linguistic Society of Japan, first as vice president from 1938 and then as president from 1967. He held this post until his death on November 14, 1971, at the age of 89. He was also known as the father of Haruhiko Kindaichi, a respected linguist in his own right.

See also: Ainu; Batchelor, John (1853–1944); Chiri, Mashiho (1909-61); Hattori, Shiro (1908–1995); History of Linguistics: Discipline of Linguistics; Japan: Language Situation; Japanese; Naert, Pierre (1916–1971).

Bibliography

Chiri M & Kindaichi K (1936). *Ainu goho gaisetsu* [*Outline of Ainu grammar*]. Tokyo: Iwanami Shoten.

Kindaichi K (1931). *Ainu jojishi yukara no kenkyu* [*Studies on the yukar*] (2 vols). Tokyo: Toyo Bunko.

Kindaichi K (1993). *Kindaichi Kyosuke zenshu* [*The collected works of Kyosuke Kindaichi*] (10 vols). Tokyo: Sanseido.

Kindaichi K (1941). *Ainu life and legends.* Tokyo: Board of tourist industry, Japanese government railways.

Kinesics

M Danesi, University of Toronto, Toronto, Canada

Introduction

The study of bodily movements, such as gestures and facial expressions, is called *kinesics* (from the Greek word *kinêsis* 'movement'). Kinesic communication is a complementary, supplementary, and at times, substitutive mode of communication to verbal (and especially vocal) language. Kinesics is a part of anthropology, psychology, and semiotics, where it is included as a topic within the general domain of nonverbal communication.

The founder of kinesics as a field of inquiry and research was the American anthropologist Ray L. Birdwhistell (1918–1994). Influenced by the views of his teacher, the great anthropologist Margaret Mead (1901–1978), and the work of David Efron (1941) in the early 1940s, Birdwhistell became interested in analyzing the way people interacted by watching films. He noticed that people seemed to transmit information unconsciously through their eye movements, facial expressions, and postures. For this reason, he came to view body language as a critical component of human interaction. His first book on the topic, *Introduction to kinesics*, was published in 1952. In it, he discussed the role of body motion and gesture in human communication.

In a series of papers and encyclopedia entries (1955, 1960, 1961, 1963, 1974, 1979), Birdwhistell defined kinesics as the study of all aspects of nonverbal communication, including gesture and touch. Today, however, some of these are studied under separate branches of semiotics (*see* **Nonverbal Communication**). Kinesic phenomena are also called 'paralanguage' within linguistics proper, indirectly

Native American languages, noting the polysynthetic, incorporating character of the Ainu verb, which is also found in Amerindian languages.

Several of Kindaichi's students went on to become prominent scholars of the Ainu language themselves. These included Shiro Hattori, a recognized dialectologist of both Hokkaido and Sakhalin Ainu, and Itsuhiko Kubodera, a compiler of Ainu oral literature. Kindaichi was also remarkable for encouraging Ainu people themselves to participate in and conduct research into their own language and oral histories. Among the most important of these scholars were Mashiho Chiri, the first Ainu to attend Tokyo Imperial University, with whom Kindaichi co-authored *Ainu goho gaisetsu* [Outline of Ainu grammar], and Shigeru Kayano, an Ainu anthropologist and politician.

Along with several native speakers of the Ainu language, notably Yukie Chiri and Matsu Kannari, Kindaichi compiled and published various yukar and other stories that might have been lost otherwise. He also helped John Batchelor's adopted daughter, Yaeko Batchelor, publish as part of her crusade against the alcoholism that was ravaging the Ainu community. Although Kindaichi believed the traditional Ainu would disappear, he resisted attempts to characterize the Ainu as a dying race, as was common at the time, and he offered instead the possibility of constructive assimilation – using his star pupil, Mashiho Chiri, as an example of Ainu genius. His work had the double effect of helping to empower a generation of Ainu scholars to become active researchers and interpreters of their culture and of providing a source of pride, in the form of the rich intellectual tradition of the yukar, for a people who for generations had been told that they had no culture or civilization to speak of.

Besides his Ainu studies, Kindaichi was active in the Linguistic Society of Japan, first as vice president from 1938 and then as president from 1967. He held this post until his death on November 14, 1971, at the age of 89. He was also known as the father of Haruhiko Kindaichi, a respected linguist in his own right.

See also: Ainu; Batchelor, John (1853–1944); Chiri, Mashiho (1909-61); Hattori, Shiro (1908–1995); History of Linguistics: Discipline of Linguistics; Japan: Language Situation; Japanese; Naert, Pierre (1916–1971).

Bibliography

Chiri M & Kindaichi K (1936). *Ainu goho gaisetsu* [*Outline of Ainu grammar*]. Tokyo: Iwanami Shoten.

Kindaichi K (1931). *Ainu jojishi yukara no kenkyu* [*Studies on the yukar*] (2 vols). Tokyo: Toyo Bunko.

Kindaichi K (1993). *Kindaichi Kyosuke zenshu* [*The collected works of Kyosuke Kindaichi*] (10 vols). Tokyo: Sanseido.

Kindaichi K (1941). *Ainu life and legends.* Tokyo: Board of tourist industry, Japanese government railways.

Kinesics

M Danesi, University of Toronto, Toronto, Canada

Introduction

The study of bodily movements, such as gestures and facial expressions, is called *kinesics* (from the Greek word *kinêsis* 'movement'). Kinesic communication is a complementary, supplementary, and at times, substitutive mode of communication to verbal (and especially vocal) language. Kinesics is a part of anthropology, psychology, and semiotics, where it is included as a topic within the general domain of nonverbal communication.

The founder of kinesics as a field of inquiry and research was the American anthropologist Ray L. Birdwhistell (1918–1994). Influenced by the views of his teacher, the great anthropologist Margaret Mead (1901–1978), and the work of David Efron (1941) in the early 1940s, Birdwhistell became interested in analyzing the way people interacted by watching films. He noticed that people seemed to transmit information unconsciously through their eye movements, facial expressions, and postures. For this reason, he came to view body language as a critical component of human interaction. His first book on the topic, *Introduction to kinesics*, was published in 1952. In it, he discussed the role of body motion and gesture in human communication.

In a series of papers and encyclopedia entries (1955, 1960, 1961, 1963, 1974, 1979), Birdwhistell defined kinesics as the study of all aspects of nonverbal communication, including gesture and touch. Today, however, some of these are studied under separate branches of semiotics (*see* **Nonverbal Communication**). Kinesic phenomena are also called 'paralanguage' within linguistics proper, indirectly

acknowledging the crucial role that they are perceived to play in speech communication.

Birdwhistell fashioned kinesic practice on the general methods and concepts of American descriptive linguistics. By collecting and analyzing relevant data on body movements, facial expressions, gestures, and the like during social interaction Birdwhistell claimed that it is possible to write a 'kinesic grammar' in the same way that linguistics write a verbal grammar by analyzing words and sentences. As he (1970: 157) put it: "The first premise in developing a notational system for body language is to assume that all movements of the body have meaning. None are accidental." He called such movements 'kinemes' (in analogy with 'phonemes'). The kineme is the prototype for a group of movements that are not identical, "but which may be used interchangeably without affecting social meaning" (Knapp, 1978: 94–95), and thus "can be construed as having a definite organization or structure, just as language is understood in terms of its grammar" (Duncan and Fiske, 1977: xi).

Because of its foundations in linguistic structuralism, kinesics was included early on within semiotics (Sebeok *et al.*, 1964; Eco, 1968). Some see it as a term encompassing the general study of nonverbal communication (Duncan, 1969: 118; Key, 1975; Harper *et al.*, 1978: 119; Knapp, 1978: 12; Raffler-Engel, 1980: 30). Others see it as neither an autonomous field of study nor a branch of semiotics; to them, it is an approach to the study of verbal behavior (Argyle, 1975; Scherer *et al.*, 1979; Scherer and Ekman, 1982).

Birdwhistell divides the study of kinesic communication into four procedures:

1. **Prekinesics**. This involves a description of the physiology of body movements (similar in nature to phonetics, which describes the physical nature of verbal sounds). These are called 'kines' (in analogy with 'phones'), constituting the minimal elements of corporeal behavior that have a differential function in a given context.
2. **Microkinesics**. This is the level of identification of 'kinemes,' the basic differential corporeal units and their variants, called 'allokines' (in analogy with 'allophones'). To make the study of kinemes systematic and standardized, Birdwhistell developed symbols for eight body sections – head, face, trunk, shoulder, hands, legs, feet, neck – called 'kinegraphs.'
3. **Social kinesics**. This is the analysis of the meaning of kinemes within social contexts.
4. **Parakinesics**. This involves identifying the modifiers of kinemes, including motion qualifiers (intensity, range, velocity of movement) and activity modifiers (mode of an entire body in motion).

Areas of Kinesic Research

Humans convey over two-thirds of their messages through the body, producing up to 700 000 physical signs, of which 1000 are different bodily postures, 5000 are hand gestures, and 250 000 are facial expressions (Morris *et al.*, 1979). The body is, in a phrase, a primary source of communication. Studying the kinesic signs that are produced with, through, or on the body has become a central objective of semiotics. Kinesic signs can be inborn (unwitting), learned (witting), or a mixture of the two. Blinking the eyes, clearing the throat, and facial flushing are innate (inborn) kinesic signals. These are often involuntary, as are, for example, facial expressions of happiness, surprise, anger, disgust, and other basic emotions. Laughing, crying, and shrugging the shoulders are examples of mixed kinesic signals. They may originate as innate actions, but cultural rules shape their timing and use. Gestures such as a wink of the eye, a thumbs up, or a military salute, are learned kinemes. The meanings of such signs vary among different cultures. Kinesic messages can give a look and feel to a conversation remembered long after spoken words fade away. They can also be structured to lie or conceal something. For example, pressing the lips together may indicate disagreement or doubt, even if the person's verbal statements convey agreement. When verbal and kinesic statements conflict, listeners will likely believe the latter more.

The main areas of research within kinesics are eye contact, posture, touch, gesture, and facial expression. Clothing is often included within the kinesic purview of research. However, as part of 'material culture,' rather than of kinesics *per se*, the meanings of clothes are usually treated under a separate branch of semiotics (*see* **Clothing: Semiotics**).

Eye Contact

Research on eye contact in various fields, from psychology to semiotics, has found that across cultures, kines such as the length of time involved in eye contact along with the pattern of contact (looking into the eyes, looking down or up, etc.) convey what kinds of social relationship people have to each other. Some patterns appear to be universal: staring is typically interpreted as a challenge; 'making eyes' at someone is normally interpreted as flirtation; narrow eyelids communicate pensiveness, and making the eyebrows come nearer together communicates thoughtfulness; raising the eyebrows conveys surprise. However, even in these there is culture-specific variation. It has been found that Southern Europeans will tend to look more into each other's eyes during conversation than

will North Americans; in some cultures males do not look into female eyes unless they are married or members of the same family. Some patterns convey specific meanings in specific cultural contexts. For example, gazing is interpreted as indicative of sexual wonder, fascination, awe, or admiration depending on the culture; staring is perceived as sexual curiosity, boldness, insolence, or stupidity – again depending on social context and culture.

A large amount of eye communication among humans unfolds in the form of unwitting signals. It has been shown, for example, that men are sexually attracted to women with large pupils, which signal unconsciously a strong and sexually tinged interest, as well as making females look younger (Sebeok, 1994). This would explain the cosmetic vogue in central Europe during the 1920s and 1930s of women using a crystalline alkaloid eye-drop liquid derived from *belladonna* ('beautiful woman' in Italian). The women of the day bought the liquid because they believed – and correctly so, it would appear – that it would enhance facial appearance and sexual attractiveness by dilating the pupils. As the linguist Karl Bühler (1934: 28) aptly observed, such signals act like regulators, eliciting or inhibiting some action or reaction.

Posture

Posture communicates a large amount of information about people's identity, gender, class, moods, motivation, and attitudes, playing a critical role in interpersonal relationships. It can be defined as a pose or movement that can be broken down into a series of specific kinetic movements conveying certain personal and/or social meanings. As in the case of eye contact, there appear to be both universals and culture-based differences in posture kinemes, as kinesic research has found. Some 'universal' posture kinemes are:

1. Slumped posture = low spirits
2. Erect posture = high spirits, energy, and confidence
3. Leaning forward posture = open and interested behavior
4. Leaning away posture = defensive or disinterested behavior
5. Crossed arms = defensive behavior
6. Uncrossed arms = willingness to listen.

Posture is prominent as a tactic in courtship displays, which may look comical or absurd to outsiders, but which constitute crucial modes of communication at key stages in the enactment of reproductive urges to the members of a society or group. 'Sexual posing' makes sense only if the appropriate physical and social contexts are present during courtship or flirtation. While such posing may be residues of ancient animal mechanisms, as some suggest, the great diversity that is evident in courtship posing across cultures suggests that it is not a mere contemporary version of instinctual mating behavior. Rather, sexual postures are shaped in large part by human notions of gender and romance and are, therefore, constantly inclined to change (Synnot, 1993). In the human species, courtship is not only a reflex of biology, but also a product of history and tradition. It is the outcome of Nature and Culture cooperating in a type of partnership that is found nowhere else in the animal realm.

Sexual posing is part of a culture's gender code that mediates peoples' perception of which bodily parts or zones are erogenous. The semiotician Michel Foucault (1926–1984) argued persuasively that 'sins of the flesh' are hardly universal (Foucault, 1976). They must be defined culturally. The Puritans of England, for instance, saw any form of sexual contact or gazing in a marriage situation as a kind of 'necessary sin.' 'Sexual temptation' is still felt by some people to be 'sinful.' On the other hand, the many 'hedonistic' rites and practices of our own and other cultures exalt and glorify the eroticism of the human body. Obviously, what is 'obscene' behavior to some is 'natural' behavior to others. While sexual urges are based in biology, perceptions of what is or is not erotic, sinful, or obscene are ensconced in cultural traditions and habits.

Touch

Another main area of kinesic research studies the social functions and manifestations of touch. The study of touch actually constitutes its own subfield, generally called 'haptics,' although it is still considered a general category of inquiry within kinesics.

A haptic pattern that reveals how research in this area is conducted is 'hand touching,' commonly called 'handshaking.' The zoologist Desmond Morris (1969) claimed that the Western form of handshaking may have started as a way to show that neither person was holding a weapon. It thus became a 'tie sign,' because of the bond it was designed to create. Throughout the centuries, the tie sign became a symbol of equality and fairness among individuals and was used to seal agreements of all kinds. Indeed, refusing to shake someone's outstretched hand was perceived, and continues to be interpreted to this day, as a 'counter-sign' of aggressiveness or as a challenge. Predictably, this form of tactile greeting reveals a high degree of cross-cultural variation. People can squeeze the hand (as Europeans and North Americans do), shake the other's hand with both hands, shake the

hand and then pat the other's back or hug him or her, lean forward or stand straight while shaking, and so on. Haptic greeting and social contact codes are not limited, of course, to manual touch patterns. Other patterns include patting someone on the arm, shoulder, or back to indicate agreement or to compliment; linking arms to indicate companionship; putting one's arm around the shoulder to indicate friendship or intimacy; holding hands with family members or a lover to express intimacy; hugging to convey happiness at seeing a friend or a family member; and so on.

Anthropologists are unclear as to why haptic codes vary so much across cultures. The reason may have a basis in culture perceptions of the body. Some people perceive the skin to be a surface 'sheath.' Others perceive the body as a 'container' or think of themselves as 'contained' in their skin. The zones of privacy that define Self-space in these cultures, therefore, include the clothes that cover the skin. Others feel instead that the Self is located down within the body shell, resulting in a totally different perception of haptic behaviors. People in such cultures are in general more tolerant of crowds, of noise levels, of the touching of hands, of eye contact, and of body odors than most North Americans are. These perceptions are bound to affect haptic behavior that underlies the formation of cultural body codes.

One aspect of haptic behavior that is shrouded in evolutionary mystery is 'lip touching' or osculation in the human species, known, of course, more commonly as 'kissing.' When the lips of adult people touch, the kissing act is perceived normally as an erotic one. But not all kissing is erotic. Kissing other parts of the face, such as the forehead, the head, and the cheeks can be a way of showing affection to children, friends, pets, etc. Erotic kissing, however, is particularly interesting as an evolutionary and cultural phenomenon, because it seems to be a kind of mock-suckling or mock-feeding action, implying vulnerability, closeness, and sensuality. However, it is not universal. It is not common in China or Japan, for instance; and it is completely unknown in many African societies. Traditional Inuit and Laplander societies are more inclined to rub noses than to kiss to engage in sexual foreplay.

Gesture

Gesture is another main area of kinesic research, although it too, like haptic communication, tends to constitute its own autonomous subfield of investigation. Gesture can be defined simply as the use of the hands, the arms, and to a lesser extent, the head to make kinesic signs. Although there are cross-cultural similarities in gesture, substantial differences also exist both in the extent to which gesture is used and in the interpretations given to its particular uses. For example, the head gestures for 'yes' and 'no' used in the Balkans seem inverted to other Europeans. In 1979, Desmond Morris, together with several of his associates at Oxford University, examined 20 gesture forms in 40 different areas of Europe. The research team found some rather fascinating things. For instance, they discovered that many of the forms had several meanings, depending on culture – a tap on the side of the head can indicate completely opposite things, 'stupidity' or 'intelligence,' according to cultural context. They also found that most of the forms were used in many countries.

Gesture is also found in primates. Chimpanzees raise their arms in the air as a signal that they want to be groomed, they stretch out their arms to beg or invite, and they have the ability to point to things (Beaken, 1996: 51). These gestures are, evidently, purposeful and regulatory of the actions of other chimps. But the number of gestural forms of which chimpanzees are capable is limited. Human gesturing, on the other hand, is productive and varied. And it can be used as a substitute for vocal language, as is the case with the many sign languages used in communities of the hearing-impaired. It can also be employed as an alternate mode of communication, as is the case with the sign languages used by religious groups during periods of imposed silence, with the hand language used by traffic personnel, and with the hand and arm language used by conductors of orchestras.

Many semioticians and linguists consider gesture to be a more fundamental form of communication than vocal language. A pragmatic confirmation of this is obtained when one doesn't speak the language of the people of a country one is visiting. In order to communicate with the people on the street, in a hotel, or in a store, one instinctively resorts to gesture in order to get a message across or to negotiate meaning. For example, if one were to describe an automobile to someone, one would instinctively use the hands to iconically portray a steering wheel and the motion used to steer a car, accompanying this gesture, perhaps, with an imitative sound of a motor. This anecdotal scenario not only suggests that gesture is a fundamental mode of communication, but also that its essentially iconic modality makes it a much more universal, and less culture-dependent, mode of message-making than vocal language.

Gesture spans the entire range of signification of vocal language (Bremer and Roodenberg, 1991; Armstrong *et al.*, 1995). Using the index finger is the most common manifestation of deixis (indicating where things are in space and time), although any

body part that can be moved directionally (lips, nose, tongue, etc.) can also be used to point out referents in the immediate environment, to indicate directions, etc. Iconic (simulative) gestures are employed commonly to represent the shape of objects: for example, to refer to a round object people the world over tend to use both hands together moving in opposite – clockwise (the right hand) and counter-clockwise (the left hand) – directions. Such a gesture is characterizable as a kind of 'spatial drawing movement.' And indeed, if the movements of the hands during gesture are transferred by some drawing instrument onto some surface such as paper, the referent of the gesture will be transferred to the paper as a visual figure. Fingers can also be used to represent symbols (by portraying the outline of the symbol). The most common use of this kind of gesturing is in the sign languages for the hearing-impaired. Lastly, symbolic gestures are often used to stand conventionally for social functions or for carrying out interactional protocols such as greeting, affirmation, negation, etc.

Some truly fascinating research by the linguist David McNeill (1992) has shown, moreover, that gesture is a complementary communicative modality of vocal language. McNeill videotaped a large number of people as they spoke, gathering a substantial amount of data on gestural behavior during vocal communication. The gestures that accompany speech are known as 'gesticulants.' McNeill's findings suggest that gesticulants are unconscious forms that convey images that cannot be shown overtly in speech, as well as images of what the speaker is thinking about. Vocal speech and gesticulation constitute a single integrated communication system in which both cooperate to express the person's intentions and thoughts.

On the basis of his findings, McNeill was able to classify gesticulants into five main categories. First, there are 'iconic' gesticulants which, as their name suggests, bear a close resemblance to the referent or referential domain of an utterance, e.g., when describing a scene from a story in which a character bends a tree back to the ground, a speaker observed by McNeill appeared to grip something and pull it back. His gesture was, in effect, a visual icon of the action talked about, revealing both his memory image and his point of view (he could have taken the part of the character or the tree, instead).

Second, there are 'metaphoric' gesticulants. These are also pictorial, but their content is abstract, rather than strictly iconic. For example, McNeill observed a male speaker announcing that he had just seen a cartoon, simultaneously raising up his hands as if offering his listener a kind of object. He was obviously not referring to the cartoon itself, but to the 'genre' of the cartoon. His gesture created and displayed this genre as if it were an object, placing it into an act of offering to the listener. This type of gesticulant typically accompanies utterances that contain expressions such as 'presenting an idea,' 'putting forth an idea,' 'offering advice,' and so on.

Third, there are 'beat' gesticulants. These resemble the beating of musical tempo. The speaker's hand moves along with the rhythmic pulsation of speech, in the form of a simple flick of the hand or fingers up and down, or back and forth. Beats are indexes, marking the introduction of new characters, summarizing the action, introducing new themes, etc. during the utterance.

Fourth, there are 'cohesive' gesticulants. These serve to show how separate parts of an utterance are supposed to hold together. Beats emphasize sequentiality, cohesives globality. Cohesives can take iconic, metaphoric, or beat form. They unfold through a repetition of the same gesticulant form, movement, or location in the gesture space. It is the repetition that is meant to convey cohesiveness.

Fifth, there are 'deictic' gesticulants – 'deixis' is the term used to designate all kinds of pointing or indicating signs. Deictic gesticulants are aimed not at an existing physical place, but at an abstract concept that had occurred earlier in the conversation. These reveal that we perceive concepts as having a physical location in space.

McNeill's work gives us a good idea as to how gesture and language complement each other in normal discourse. As Frutiger (1989: 112) has also observed, accompanying gesture forms reveal an inner need to support what one is saying orally: "If on a beach, for example, we can hardly resist drawing with the finger on the smooth surface of the sand as a means of clarifying what we are talking about." McNeill's gesticulant categories are actually subtypes of the more generic category of gesture called an 'illustrator' by kinesics researchers. Other categories are emblems, affect displays, regulators, and adaptors:

1. **Illustrators**: As just discussed, these literally illustrate vocal utterances. Examples are the circular hand movements when talking of a circle, moving the hands far apart when talking of something large, moving both the head and hands in an upward direction when saying *Let's go up*.
2. **Emblems**: These directly translate words or phrases. Examples are the *Okay* sign, the *Come here* sign, the hitchhiking sign, waving, and obscene gestures.
3. **Affect Displays**: These communicate emotional meaning. Examples are the typical hand movements that accompany states and expressions of

happiness, surprise, fear, anger, sadness, contempt, disgust, etc.

4. **Regulators**: These monitor, maintain, or control the speech of someone else. Examples include the hand movements for *Keep going, Slow down, What else happened?*
5. **Adaptors**: These are used to satisfy some need. Examples include scratching one's head when puzzled, rubbing one's forehead when worried, and so on.

Some societies have developed complementary 'gesture languages.' These share structural and semantic features with vocal languages. One of the best-known examples is the hand language developed by the Plains people of North America as a means of communication between tribes with different vocal languages. The manual signs represent things in nature, ideas, emotions, and sensations. For example, the sign for a white person is made by drawing the fingers across the forehead, indicating a hat. Special signs exist also for each tribe and for particular rivers, mountains, and other natural features. The sensation of cold is indicated by a shivering motion of the hands in front of the body, and the same sign is used for 'winter' and for 'year,' because the Plains peoples count years in terms of winters. Slowly turning the hand, relaxed at the wrist, means vacillation, doubt, or possibility; a modification of this sign, with a quicker movement, is the question sign. The hand language is so elaborate that a detailed conversation is possible using the gestures alone (Mallery, 1972).

Gestures may also be used for sacred symbolic purposes, revealing the ancientness and thus 'mythic symbolism,' associated with hand communication. For example, in Christianity the 'sign of the cross' is a gesture that aims to recreate the central event of Christianity – the Crucifixion. In Buddhism, the gestures known as *Mudras* are used during ceremonies to represent meditation, reasoning, doctrine, protection and request, enlightenment, unification of matter, and spirit. The 'devil's hand,' with the index and little finger raised, on the other hand, belongs to the domain of superstition, symbolizing, in some cultures, a horned figure intended to ward off the evil eye and in others a sign of 'cuckoldry.'

Facial Expression

Probably inspired by general work in kinesics, in 1963, psychologist Paul Ekman established the Human Interaction Laboratory in the Department of Psychiatry at the University of California at San Francisco for the purpose of studying one particular aspect of kinesic behavior – facial expression. He was joined by Wallace V. Friesen in 1965 and Maureen O'Sullivan in 1974. Over the years, Ekman and his team have identified, for example, certain facial expressions as universal signs of specific emotions (e.g., Ekman, 1980, 1982, 1985, 2003; Ekman and Friesen, 1975). They have shown this by breaking down facial expressions into characteristic components – eyebrow position, eye shape, mouth shape, nostril size, etc. – which in various combinations determine the form and meaning of the expression. Ekman and his team have, over the years, found very little variation across cultures in the nature of the components and in their combinations. Indeed, they have shown that it is possible to write a 'grammar' of the face that shows less cross-cultural variation than do language grammars (Peck, 1987).

Conclusion

Kinesics has become an important branch of semiotics because it documents the power of nonverbal signs in regulating how people behave in certain social situations. Such signs are products of both our biological ancestry and of specific cultural histories. The sounds made by the body (sneezing, coughing, burping, etc.), the fluids that issue forth from it, the eye patterns that are used when speaking, the kinds of touching routines humans utilize during discourse, the facial expressions utilized in conveying feelings, and the kinds of gestures and gesticulations that are found throughout the world are all interpreted semiotically as signifiers of both universal and culture-specific meanings. These are the outcome of Nature and Culture cooperating in an exclusive form of partnership.

Discoveries in neuroscience have shown that nonverbal signs are produced and processed differently from words. Spoken language is processed in the cerebral cortex, a more developed area of the brain that is unique to human beings. In contrast, nonverbal cues – such as smiling, staring, and clenching the fists – are processed in lower, more primitive areas such as the limbic system. People often produce and receive nonverbal cues without conscious awareness of doing so.

Because kinesics has produced a substantial and substantive literature, it is now not uncommon to separate the various areas mentioned in this entry into separate subfields. For example, gesture is now studied separately. Similarly, the study of facial expression and touching patterns are treated as separate areas of investigation.

See also: Body Language; Clothing: Semiotics; Dance; Nonverbal Communication; Proxemics; Sign Language: Semiotic Approaches.

Bibliography

Argyle M (1975). *Bodily communication.* London: Methuen.

Armstrong D F, Stokoe W C & Wilcox S E (1995). *Gesture and the nature of language.* Cambridge: Cambridge University Press.

Beaken M (1996). *The making of language.* Edinburgh: Edinburgh University Press.

Birdwhistell R L (1952). *Introduction to kinesics.* Louisville, KY: University of Louisville Press.

Birdwhistell R L (1955). 'Background to kinesics.' *ETC 13,* 10–18.

Birdwhistell R L (1960). 'Kinesics and communication.' In Carpenter E & McLuhan M (eds.) *Explorations in communication.* New York: Beacon. 54–64.

Birdwhistell R L (1961). 'Paralanguage 25 years after Sapir.' In Laver J & Hutcheson S (eds.) *Communication in face to face interaction.* Harmondsworth: Penguin. 82–100.

Birdwhistell R L (1963). 'The kinesic level in the investigation of the emotions.' In Knapp P (ed.) *Symposium on expressions of the emotions in man.* New York: International University Press. 123–139.

Birdwhistell R L (1974). 'The language of the body.' In Silverstein A (ed.) *Human communication.* Hillsdale, NJ: Erlbaum. 203–211.

Birdwhistell R L (1979). 'Kinesics.' In *International encyclopedia of the social sciences,* vol. 8. New York: Macmillan. 379–385.

Bremer J & Roodenburg H (eds.) (1991). *A cultural history of gesture.* Ithaca: Cornell University Press.

Bühler K (1934). *Sprachtheorie: Die Darstellungsfunktion der Sprache.* Jena: Fischer.

Duncan S (1969). 'Nonverbal communication.' *Psychological Bulletin 72,* 118–137.

Duncan S & Fiske D W (1977). *Face-to-face interaction.* Hillsdale, NJ: Erlbaum.

Eco U (1968). *Einführung in die Semiotik.* München: Fink.

Efron D (1941). *Gesture, race, and culture.* The Hague: Mouton.

Ekman P (1976). 'Movements with precise meanings.' *Journal of Communication 26,* 14–26.

Ekman P (1980). 'The Classes of Nonverbal Behavior.' In Raffler-Engel W (ed.) *Aspects of nonverbal communication.* Lisse: Swets and Zeitlinger. 89–102.

Ekman P (1982). 'Methods for measuring facial action.' In Scherer K R & Ekman P (eds.) *Handbook of methods in nonverbal behavior.* Cambridge: Cambridge University Press. 45–90.

Ekman P (1985). *Telling lies.* New York: Norton.

Ekman P (2003). *Emotions revealed.* New York: Holt.

Ekman P & Friesen W (1975). *Unmasking the face.* Englewood Cliffs: Prentice-Hall.

Foucault M (1976). *The history of sexuality* (vol. 1). London: Allen Lane.

Frutiger A (1989). *Signs and symbols.* New York: Van Nostrand.

Harper R G *et al.* (1978). *Nonverbal communication: the state of the art.* New York: John Wiley.

Key M R (1975). *Paralanguage and kinesics.* Metuchen, NJ: Scarecrow.

Knapp M L (1978). *Nonverbal communication in human interaction.* New York: Holt.

Mallery G (1972). *Sign language among North American Indians compared with that among other peoples and deaf-mutes.* The Hague: Mouton.

McNeill D (1992). *Hand and mind: what gestures reveal about thought.* Chicago: University of Chicago Press.

Morris D (1969). *The human zoo.* London: Cape.

Morris D *et al.* (1979). *Gestures: Their origins and distributions.* London: Cape.

Peck S R (1987). *Atlas of facial expression.* Oxford: Oxford University Press.

Raffler-Engel W von (ed.) (1980). *Aspects of nonverbal communication.* Lisse: Swets.

Scherer K R & Ekman P (eds.) (1982). *Handbook of methods in nonverbal behavior research.* Cambridge: Cambridge University Press.

Scherer K R, Wallbott H G & Scherer U (1979). 'Methoden zur Klassifikation von Bewegungsverhalten: Ein funktionaler Ansatz.' *Zeitschrift für Semiotik 1,* 177–192.

Sebeok T A (1994). *Signs: an introduction to semiotics.* Toronto: University of Toronto Press.

Sebeok T A *et al.* (eds.) (1964). *Approaches to semiotics.* The Hague: Mouton.

Synnott A (1993). *The body social: symbolism, self and society.* London: Routledge.

Kingman Reef: Language Situation

Editorial Team

Wet or awash most of the time, with a maximum elevation of approximately 1 m, Kingman Reef is situated in the North Pacific Ocean, about half way between Hawaii and American Samoa. An unincorporated territory of the United States, Kingman Reef is administered from Washington by the Fish and Wildlife Service of the Department of the Interior. It is completely uninhabited, but marine flora and fauna thrive there. Kingman Reef was used for a brief period in the 1930s as a halfway station for Pan American Airways flying boats.

Bibliography

Argyle M (1975). *Bodily communication.* London: Methuen.

Armstrong D F, Stokoe W C & Wilcox S E (1995). *Gesture and the nature of language.* Cambridge: Cambridge University Press.

Beaken M (1996). *The making of language.* Edinburgh: Edinburgh University Press.

Birdwhistell R L (1952). *Introduction to kinesics.* Louisville, KY: University of Louisville Press.

Birdwhistell R L (1955). 'Background to kinesics.' *ETC 13,* 10–18.

Birdwhistell R L (1960). 'Kinesics and communication.' In Carpenter E & McLuhan M (eds.) *Explorations in communication.* New York: Beacon. 54–64.

Birdwhistell R L (1961). 'Paralanguage 25 years after Sapir.' In Laver J & Hutcheson S (eds.) *Communication in face to face interaction.* Harmondsworth: Penguin. 82–100.

Birdwhistell R L (1963). 'The kinesic level in the investigation of the emotions.' In Knapp P (ed.) *Symposium on expressions of the emotions in man.* New York: International University Press. 123–139.

Birdwhistell R L (1974). 'The language of the body.' In Silverstein A (ed.) *Human communication.* Hillsdale, NJ: Erlbaum. 203–211.

Birdwhistell R L (1979). 'Kinesics.' In *International encyclopedia of the social sciences,* vol. 8. New York: Macmillan. 379–385.

Bremer J & Roodenburg H (eds.) (1991). *A cultural history of gesture.* Ithaca: Cornell University Press.

Bühler K (1934). *Sprachtheorie: Die Darstellungsfunktion der Sprache.* Jena: Fischer.

Duncan S (1969). 'Nonverbal communication.' *Psychological Bulletin 72,* 118–137.

Duncan S & Fiske D W (1977). *Face-to-face interaction.* Hillsdale, NJ: Erlbaum.

Eco U (1968). *Einführung in die Semiotik.* München: Fink.

Efron D (1941). *Gesture, race, and culture.* The Hague: Mouton.

Ekman P (1976). 'Movements with precise meanings.' *Journal of Communication 26,* 14–26.

Ekman P (1980). 'The Classes of Nonverbal Behavior.' In Raffler-Engel W (ed.) *Aspects of nonverbal communication.* Lisse: Swets and Zeitlinger. 89–102.

Ekman P (1982). 'Methods for measuring facial action.' In Scherer K R & Ekman P (eds.) *Handbook of methods in nonverbal behavior.* Cambridge: Cambridge University Press. 45–90.

Ekman P (1985). *Telling lies.* New York: Norton.

Ekman P (2003). *Emotions revealed.* New York: Holt.

Ekman P & Friesen W (1975). *Unmasking the face.* Englewood Cliffs: Prentice-Hall.

Foucault M (1976). *The history of sexuality* (vol. 1). London: Allen Lane.

Frutiger A (1989). *Signs and symbols.* New York: Van Nostrand.

Harper R G *et al.* (1978). *Nonverbal communication: the state of the art.* New York: John Wiley.

Key M R (1975). *Paralanguage and kinesics.* Metuchen, NJ: Scarecrow.

Knapp M L (1978). *Nonverbal communication in human interaction.* New York: Holt.

Mallery G (1972). *Sign language among North American Indians compared with that among other peoples and deaf-mutes.* The Hague: Mouton.

McNeill D (1992). *Hand and mind: what gestures reveal about thought.* Chicago: University of Chicago Press.

Morris D (1969). *The human zoo.* London: Cape.

Morris D *et al.* (1979). *Gestures: Their origins and distributions.* London: Cape.

Peck S R (1987). *Atlas of facial expression.* Oxford: Oxford University Press.

Raffler-Engel W von (ed.) (1980). *Aspects of nonverbal communication.* Lisse: Swets.

Scherer K R & Ekman P (eds.) (1982). *Handbook of methods in nonverbal behavior research.* Cambridge: Cambridge University Press.

Scherer K R, Wallbott H G & Scherer U (1979). 'Methoden zur Klassifikation von Bewegungsverhalten: Ein funktionaler Ansatz.' *Zeitschrift für Semiotik 1,* 177–192.

Sebeok T A (1994). *Signs: an introduction to semiotics.* Toronto: University of Toronto Press.

Sebeok T A *et al.* (eds.) (1964). *Approaches to semiotics.* The Hague: Mouton.

Synnott A (1993). *The body social: symbolism, self and society.* London: Routledge.

Kingman Reef: Language Situation

Editorial Team

Wet or awash most of the time, with a maximum elevation of approximately 1 m, Kingman Reef is situated in the North Pacific Ocean, about half way between Hawaii and American Samoa. An unincorporated territory of the United States, Kingman Reef is administered from Washington by the Fish and Wildlife Service of the Department of the Interior. It is completely uninhabited, but marine flora and fauna thrive there. Kingman Reef was used for a brief period in the 1930s as a halfway station for Pan American Airways flying boats.

Kinship Expressions and Terms

J G Heath, University of Michigan, Ann Arbor, MI, USA

All languages have expressions that can be glossed 'father,' 'wife,' 'brother-in-law,' etc., covering consanguineal, spousal, and affinal categories. However, from the beginning of serious cross-cultural study, it has been apparent that societies differ in the kinship categories used. To facilitate comparison, an 'etic grid' is necessary; namely, an open-ended universe of kin-types, such as 'father's sister,' corresponding to the denotata of the infinitely expansible possessive construction seemingly found in all languages. The minimal nuclear components are parent (Pa), child (Ch), and spouse (Sp), a combinatory algorithm (PaPaSp = parent's parent's spouse), and cross-cutting features, including sex of referent (Ch + female = daughter), sex of linking relative (male linking relative + Ch + female = man's daughter), and relative age within a generation (PaSb + senior = senior sibling). The initial linking relative (e.g., *you* in *your father*) is referred to as 'ego' or 'propositus.'

In practice, most anthropologists and linguists operate with atomic categories based on a prototypical set of flesh-and-blood nuclear family members: Fa(ther), Mo(ther), Br(other), Si(ster), So(n), Da(ughter), Wi(fe), and Hu(sband). Combinations are then of the type FaMoBrWi (father's mother's brother's wife).

This grid permits analysis of the denotational scope of each kin term and, ultimately the entire system, by identifying the complete set of kin-types onto which the term maps. Some terms, such as English *mother* and *brother*, denote unique kin-types. Others, such as English *uncle*, correspond semantically to finite sets of kin-types. However, in languages with 'classificatory' kin-term systems, each such term maps onto an unbounded set of kin-types. With such languages, a common approach is to generate the more complex kin-types from simpler ones by expansion rules that link kin-types closer to propositus, with kin-types in the denoted set farther away; doing so entails converse reduction rules that go from complex to simple types. If in a system the term for MoBrDa expands to include MoMoBrDaDa, then conversely, MoMoBrDaDa reduces terminologically to MoBrDa; they are both denoted by the same term in such a system of terms.

Kinship Systems

Ideally, one can elicit the kinship system of a previously unstudied language by sitting down with an informant for a few hours, presenting a reasonable range of kin-types (perhaps using the informant's actual kin), and recording the corresponding kin terms. A great milestone in kinship study, still widely mined for data, was Morgan's (1870) analytical compendium of term/kin-type mappings from American Indian languages based on detailed questionnaires filled out by local correspondents (Trautmann, 1987). In interdisciplinary museum expeditions, such as those led by Alfred Haddon in the South Pacific in the first decades of the 20th century, it was common to include an anthropologist who would rapidly record the 'social structure' – that is, the kinship terminology – of each tribe visited in this fashion.

The resulting systems were taken as reflecting synchronic native models of how the societies worked, with particular reference to matriliny versus partriliny and to such issues as bigamous marriage involving sibling groups. However, because kin-term systems were often mismatched with observable social practices, ethnologists inferred that the kin terms could reflect social structures from an earlier age that were preserved in linguistic formaldehyde. Evolutionary models were later dismissed in favor of synchronic interpretations, first by early functional-structuralists (e.g., Radcliffe-Brown, 1924, 1950), and later by high structuralists (Lévi-Strauss, 1949) under the influence of linguistic structuralism. In recent decades, anthropology has become heavily concerned with historical processes, which has been reflected in a revival of interest in the evolution of kinship systems (Dole, 1972).

As an example of kinship typology, consider the three male kin-types: Fa, FaBr, and MoBr. The following four groupings are attested: (1) all three terminologically distinct (Sudanese systems); (2) Fa versus {FaBr, MoBr} (Eskimo systems, as in English), emphasizing the distinction between direct and collateral lines; (3) {Fa, FaBr} versus MoBr (Iroquois systems), known technically as bifurcate merging, merging same-sex siblings; and (4) all three merged into one category (Hawaiian systems), emphasizing generation. A hypothetical fifth system with (Fa, MoBr) terminologically distinguished from FaBr is unknown. The treatment of female kin-types Mo, MoSi, and FaSi is almost always parallel within a given system type; for example, Iroquois systems distinguish {Mo, MoSi} from FaSi. Usually (but not always), parental-generation kin-types that are merged terminologically as referents are also undifferentiated as linking relatives. For example, the child of an *uncle* (whether MoBr or FaBr) is *cousin* in English.

Two initially baffling systems show the same (bifurcate merging) cuts as Iroquois among the six primary parental-generation kin-types, but skew their extensions to lower generations. Crow systems have a matrilineal skewing rule by which SiSo or (woman's) So is terminologically merged with Fa; note that (ego's) Fa and SiSo are in the same matrilineal descent group, which differs from ego's own. Omaha systems have a structurally similar patrilineal skewing rule; for example, merging MoBrDa with Mo, both of whom are in ego's Mo's patrilineal descent group.

Meaning of Kin Terms

One is naturally drawn to seek explanations for this cross-linguistic variation in the form of correlation with cross-cultural differences in social practices. However, doing so turns out to be more difficult than it seems. To begin with, there are at least four social parameters that might be considered: (1) society-wide descent groups (matrilineal or patrilineal), including binary divisions (moieties); (2) marriage rules; (3) postnuptial residence (e.g., with the father of the bride); and (4) inheritance (e.g., from father to son). Unfortunately, these social phenomena operate at different levels and are often nuanced and negotiable, rather than binaristic and rigid. Although the parameters may sometimes converge (e.g., patrilineal descent, patrilocal residence, father-to-son inheritance), there are many cases where they do not. The structural-functional interpretation of "a kinship system as a working system linking human beings together in an orderly arrangement of interactions" (Radcliffe-Brown, 1950) has been plagued by evidence that certain kin terms lump together kin-types associated with quite different behaviors and affective relations.

Yet, it seems reasonable to suppose that kin-term semantics and prevailing sociocultural practices are not unrelated. For example, the fact that Americans have unique terms for parents, but lump the siblings of these parents (and even the spouses of these siblings) together as *aunt* and *uncle*, correlates with the autonomy of the nuclear family in this society. Likewise, societies that have well-developed descent systems and, say, a formalized levirate by which orphaned children are adopted by a brother of the deceased father, are good candidates for systems with Fa = FaBr (distinct from MoBr).

The suggestive but imperfect correlations between terminology and sociocultural practices have led to deep divisions among anthropologists as to what kin terms 'mean.' At one pole are exponents of formal kinship semantics, who argue that kin terms constitute autonomous formal systems, decomposable into nuclear primitives and expansion/reduction rules as sketched above. One of the founders of this approach, Floyd Lounsbury, demonstrated that this generative model provided the first accurate analysis of Morgan's Crow- and Omaha-type data, permitting the identification of previously unsuspected parametric variants within both (1956, 1964a, 1964b). He showed that previous attempts to explain these systems using sociocultural correlations were empirically inaccurate beyond the simplest kin-types. Lounsbury's model was later extended by Harold Scheffler to Australian Aboriginal and other difficult systems (1971, 1978).

The modularity of this approach was countered by other anthropologists who argued that sociocultural contextualization was needed. The most outspoken critic, David Schneider (1968, 1984), questioned the core assumption of the formalists; that is, universal concepts of marriage and procreation (or conception) as the foundation of nuclear basic kin-types (for assessments, see Feinberg and Ottenheimer, 2001). For Schneider, kinship systems are based on subjective and therefore culture-specific concepts; for example, Americans have a distinctive notion of blood at the center of their kinship system. Schneider and others emphasized the ways different cultures recognize paternity (and even maternity), citing the widespread nonbiological uses of kinship concepts (e.g., *father* as applied to God, to George Washington, or to priests), and arguing that anthropologists must discover these concepts through sensitive participant-observation. Similar attacks on universal 'marriage' were launched by British anthropologists (Needham, 1971).

Ethnographic complications include spirit-child conception, god-parents, compadres and comadres, co-spouses, ex-spouses, adoption, step-kin, namesake relationships, formal renunciation of genealogical relationships, and tightly knit brotherhoods and sisterhoods. Recent phenomena (gay marriage, artificial insemination, test-tube babies, surrogate mothers, and cloning) add further layers of complexity.

Cognitive Semantics

Cognitive anthropologists and linguists have grappled with tensions between extralinguistic conceptual structure and lexical categories (Miller and Johnson-Laird, 1976), and have of course reexamined kinship among other domains. Prototype theory has been influential, but no clear consensus on kinship semantics has emerged.

Hirshfeld (1984) argued that a formalist model based on genealogical primitives is inconsistent with

cognitive theories of meaning. He dismissed the claim that nongenealogical uses of kin terms are (experienced by natives as) metaphorical and pointed out that the time-bound acts of insemination and childbearing fail to account for the enduring nature of kinship bonds. He suggested that normative attributions of parent-child physical resemblance play a key role.

Anna Wierzbicka's theoretical program emphasized a small number of universal lexical primitives, from which the meaning of other lexical items is derived. Addressing kinship (1992), she identified MOTHER, FATHER, HUSBAND, and WIFE as primitives. She partially accepted the argument that the two parental categories are universal, based on cross-cultural recognition of paternal as well as maternal roles in procreation.

Read (1984) argued that American kinship terminology can be modeled as a modular algebraic system independent of its linkage to concrete kin categories (such as kin-types).

Linguistic Studies

Anthropologists have long been aware of terminological reductions and shifts from referential to vocative forms of kinship, generally interpreting them as pragmatic skewings of vocatives. Linguists have now brought out much deeper language-internal complexity. There may be several coexisting subsets of nouns: referential terms (*father*), vocatives (*hey Dad!*), deceased kin (*my late uncle*), bereaved kin (*orphan*), terms for nonhumans, and dyadic derivatives (*mother*-Dyad = a mother and her child). In some languages, these subsystems have their own suppletive terms and/or make different semantic cuts. For example, English *orphan*, *widow*, and *widower* are a defective set of bereaved kin terms; some languages have additional terms with glosses like 'bereaved MoBr.' Tubatulabal (California) has a suffix roughly glossable 'last surviving X' (X = e.g., Br; Voegelin, 1935).

Broad kinship relations among two or more referents may be encoded in nonsingular pronouns. Dalabon (Australia) has two forms for each dual pronominal, as in 'they(Dual)-went,' depending on whether the two referents are 'harmonic' based primarily on generation sets (Alpher, 1982; cf. Koch, 1982 on Kaytej).

One semantic type of kinship verb is glossable 'X calls Y "mother"' (cf. *Don't "mother" me!*). Another is glossable 'X mothers Y,' where (depending on the language) X is any of Y's (classificatory) mothers, X is Y's actual (prototypical) mother, or X behaves like a mother to Y. Sometimes at least certain 'referential kin terms' are morphological verbs or relative clauses (Amith and Smith-Stark, 1994; Evans, 2000). Even in other languages, the precise semantics of such verbs as 'bear/sire (child)' clarifies the conceptual foundations of kinship.

Whenever speaker S and addressee A are both related to a referent R, there is a choice between egocentric ('my . . .') and altercentric ('your . . .') perspectives. Actual usage may be controlled by pragmatic principles, but these can be intricate (Merlan, 1982). The forced choice between the two perspectives is obviated in some Australian languages in which 'triangular' kin terms simultaneously specify S's and A's relationships to R (Laughren, 1982; Merlan, 1989).

Special registers can be associated with speech to (or in the presence of) particular kin: bawdy registers for 'joking relationships' or respectful registers used in the presence of senior affines. These registers may have extensive vocabularies, including suppletive kin terms, thereby usually making a reduced set of distinctions (Dixon, 1971). The same applies for registers associated with male initiation rites (Hale, 1982).

Kinship categories may be associated with body parts. In this case, touching the body part may replace the spoken term, in isolation, or timed to fill a 'slot' in an utterance (Heath, 1982). These kin/body associations are metonymic, alluding to stereotypical behaviors (carrying in the womb, carrying piggyback, sleeping thigh to thigh, averting one's eye). The fact that carrying in the womb is treated as just one among several bodily interactions suggests that the sharp distinction between 'genealogical/biological' and 'affective/behavioral' levels of meaning in anthropological kinship studies may be misguided.

See also: Cognitive Anthropology; Identity in Sociocultural Anthropology and Language.

Bibliography

Alpher B (1982). 'Dalabon dual-subject prefixes, kinship categories, and generation skewing.' In Heath J *et al.* (eds.). 19–30.

Amith J D & Smith-Stark T C (1994). 'Predicate nominal and transitive verbal expressions of interpersonal relations.' *Linguistics 32*, 511–547.

Dixon R M W (1971). 'A method of semantic description.' In Steinberg D & Jakobovitz L (eds.) *Semantics; an interdisciplinary reader in philosophy, linguistics, and psychology.* Cambridge: Cambridge University Press. 436–471.

Dole G E (1972). 'Developmental sequences of kinship patterns.' In Reining P (ed.) *Kinship studies in the*

Morgan centennial year. Washington, DC: Anthropological Society of Washington. 134–166.
Evans N (2000). 'Kinship verbs.' In Vogel P M & Comrie B (eds.) *Approaches to the typology of word classes*. Berlin: Mouton de Gruyter. 103–172.
Feinberg R & Ottenheimer M (eds.) (2000). *The cultural analysis of kinship: the legacy of David M. Schneider.* Urbana, IL: University of Illinois Press.
Godelier M, Trautmann T R & Tjon Sie Fat F E (eds.) (1998). *Transformations of kinship*. Washington: Smithsonian Institution Press.
Hale K (1982). 'The logic of Damin kinship terminology.' In Heath J *et al.* (eds.). 31–37.
Heath J (1982). 'Where is that (knee)?: basic and supplementary kin terms in Dhuwal (Yuulngu/Murngin).' In Heath J *et al.* (eds.). 40–63.
Heath J, Merlan F & Rumsey A (eds.) (1982). *Languages of kinship in Aboriginal Australia* (Oceania Linguistic Monographs, 24.). Sydney: University of Sydney.
Hirschfeld L (1984). 'Kinship and cognition.' *Current Anthropology 2,* 217–242.
Koch H J (1982). 'Kinship categories in Kaytej pronouns.' In Heath J *et al.* (eds.). 64–71.
Laughren M (1982). 'Warlpiri kinship structure.' In Heath J *et al.* (eds.). 72–85.
Levi-Strauss C (1949). *Les structures élémentaires de la parenté*. Paris: Presses Universitaires de France.
Lounsbury F G (1956). 'A semantic analysis of Pawnee kinship usage.' *Language 32,* 158–194.
Lounsbury F G (1964a). 'A formal account of the Crow- and Omaha-type kinship terminologies.' In Goodenough W H (ed.) *Explorations in cultural anthropology.* New York: McGraw-Hill. 351–393.
Lounsbury F G (1964b). 'The structural analysis of kinship semantics.' In Lunt H G (ed.) *Proceedings of the Ninth International Congress of Linguistics.* The Hague: Mouton. 1073–1093.
Merlan F (1982). '"Egocentric" and "altercentric" usage of kin terms in Manarayi.' In Heath J *et al.* (eds.). 125–140.
Merlan F (1989). 'Jawoyn relationship terms: interactional dimensions of Australian kin classification.' *Anthropological Linguistics 31,* 27–63.
Miller G A & Johnson-Laird P N (1976). *Language and perception.* Cambridge MA: Harvard University Press.
Morgan L H (1870). *Systems of consanguinity and affinity of the human family.* Washington, DC: Smithsonian.
Needham R (1971). 'Remarks on the analysis of kinship and marriage.' In Needham (ed.) *Rethinking kinship and marriage.* London: Tavistock. 1–34.
Radcliffe-Brown A R (1924). 'The mother's brother in South Africa.' *South African Journal of Science 21, 542–555.*
Radcliffe-Brown A R (1950). 'Introduction.' In Radcliffe-Brown A R (ed.) *African systems of kinship and marriage.* London: Oxford University Press. 1–85.
Read D W (1984). 'An algebraic account of the American kinship terminology.' *Current Anthropology 25,* 417–440.
Scheffler H W (1971). *A study in structural semantics: the Siriono kinship system.* Englewood Cliffs, NJ: Prentice-Hall.
Scheffler H W (1978). *Australian kin classification.* Cambridge: Cambridge University Press.
Schneider D M (1968). *American kinship: a cultural account.* Englewood Cliffs, NJ: Prentice-Hall.
Schneider D M (1984). *A critique of the study of kinship.* Ann Arbor, MI: University of Michigan Press.
Trautmann T R (1987). *Lewis Henry Morgan and the invention of kinship.* Berkeley: University of California Press.
Voegelin C F (1935). 'Tübatulabal grammar.' *University of California Publications in American Archaeology and Ethnology 34(2),* 55–190.
Wierzbicka A (1992). *Semantics, culture, and cognition.* New York: Oxford University Press.

Kinyarwanda

A Kimenyi, California State University at Sacramento, Sacramento, CA, USA

General Background

Kinyarwanda, the national language of Rwanda is probably, after Swahili the second largest spoken language in the Bantu group. It is a sister dialect of Kirundi, the national language of Burundi, and Giha, another dialect spoken in Tanzania. Despite the genocide that took place, taking the lives of more than 1 million Tutsi, it is spoken by perhaps more than 20 million people. Rwanda has approximately 9 million people right now, Burundi has approximately 7 million, but besides the Giha speakers there are also ethnic Banyarwanda in Southern Uganda in the Kigezi district known as Bafumbira. Other Kinyarwanda speakers are Banyamulenge in Southern Kivu, ethnic Banyarwanda in Masisi, and Rutshuro in Northern Kivu in the Democratic Republic of Congo. Kinyarwanda belongs to the interlacustrine (Great Lakes) Bantu languages.

Writing System

Although Kinyarwanda has both long vowels and short vowels as well as high tones or no tones on syllables, the official orthography does not mark

Morgan centennial year. Washington, DC: Anthropological Society of Washington. 134–166.
Evans N (2000). 'Kinship verbs.' In Vogel P M & Comrie B (eds.) *Approaches to the typology of word classes.* Berlin: Mouton de Gruyter. 103–172.
Feinberg R & Ottenheimer M (eds.) (2000). *The cultural analysis of kinship: the legacy of David M. Schneider.* Urbana, IL: University of Illinois Press.
Godelier M, Trautmann T R & Tjon Sie Fat F E (eds.) (1998). *Transformations of kinship*. Washington: Smithsonian Institution Press.
Hale K (1982). 'The logic of Damin kinship terminology.' In Heath J *et al.* (eds.). 31–37.
Heath J (1982). 'Where is that (knee)?: basic and supplementary kin terms in Dhuwal (Yuulngu/Murngin).' In Heath J *et al.* (eds.). 40–63.
Heath J, Merlan F & Rumsey A (eds.) (1982). *Languages of kinship in Aboriginal Australia* (Oceania Linguistic Monographs, 24.). Sydney: University of Sydney.
Hirschfeld L (1984). 'Kinship and cognition.' *Current Anthropology 2,* 217–242.
Koch H J (1982). 'Kinship categories in Kaytej pronouns.' In Heath J *et al.* (eds.). 64–71.
Laughren M (1982). 'Warlpiri kinship structure.' In Heath J *et al.* (eds.). 72–85.
Levi-Strauss C (1949). *Les structures élémentaires de la parenté*. Paris: Presses Universitaires de France.
Lounsbury F G (1956). 'A semantic analysis of Pawnee kinship usage.' *Language 32,* 158–194.
Lounsbury F G (1964a). 'A formal account of the Crow- and Omaha-type kinship terminologies.' In Goodenough W H (ed.) *Explorations in cultural anthropology.* New York: McGraw-Hill. 351–393.
Lounsbury F G (1964b). 'The structural analysis of kinship semantics.' In Lunt H G (ed.) *Proceedings of the Ninth International Congress of Linguistics.* The Hague: Mouton. 1073–1093.
Merlan F (1982). '"Egocentric" and "altercentric" usage of kin terms in Manarayi.' In Heath J *et al.* (eds.). 125–140.
Merlan F (1989). 'Jawoyn relationship terms: interactional dimensions of Australian kin classification.' *Anthropological Linguistics 31,* 27–63.
Miller G A & Johnson-Laird P N (1976). *Language and perception.* Cambridge MA: Harvard University Press.
Morgan L H (1870). *Systems of consanguinity and affinity of the human family.* Washington, DC: Smithsonian.
Needham R (1971). 'Remarks on the analysis of kinship and marriage.' In Needham (ed.) *Rethinking kinship and marriage.* London: Tavistock. 1–34.
Radcliffe-Brown A R (1924). 'The mother's brother in South Africa.' *South African Journal of Science 21,* 542–555.
Radcliffe-Brown A R (1950). 'Introduction.' In Radcliffe-Brown A R (ed.) *African systems of kinship and marriage.* London: Oxford University Press. 1–85.
Read D W (1984). 'An algebraic account of the American kinship terminology.' *Current Anthropology 25,* 417–440.
Scheffler H W (1971). *A study in structural semantics: the Siriono kinship system.* Englewood Cliffs, NJ: Prentice-Hall.
Scheffler H W (1978). *Australian kin classification.* Cambridge: Cambridge University Press.
Schneider D M (1968). *American kinship: a cultural account.* Englewood Cliffs, NJ: Prentice-Hall.
Schneider D M (1984). *A critique of the study of kinship.* Ann Arbor, MI: University of Michigan Press.
Trautmann T R (1987). *Lewis Henry Morgan and the invention of kinship.* Berkeley: University of California Press.
Voegelin C F (1935). 'Tübatulabal grammar.' *University of California Publications in American Archaeology and Ethnology 34(2),* 55–190.
Wierzbicka A (1992). *Semantics, culture, and cognition.* New York: Oxford University Press.

Kinyarwanda

A Kimenyi, California State University at Sacramento, Sacramento, CA, USA

General Background

Kinyarwanda, the national language of Rwanda is probably, after Swahili the second largest spoken language in the Bantu group. It is a sister dialect of Kirundi, the national language of Burundi, and Giha, another dialect spoken in Tanzania. Despite the genocide that took place, taking the lives of more than 1 million Tutsi, it is spoken by perhaps more than 20 million people. Rwanda has approximately 9 million people right now, Burundi has approximately 7 million, but besides the Giha speakers there are also ethnic Banyarwanda in Southern Uganda in the Kigezi district known as Bafumbira. Other Kinyarwanda speakers are Banyamulenge in Southern Kivu, ethnic Banyarwanda in Masisi, and Rutshuro in Northern Kivu in the Democratic Republic of Congo. Kinyarwanda belongs to the interlacustrine (Great Lakes) Bantu languages.

Writing System

Although Kinyarwanda has both long vowels and short vowels as well as high tones or no tones on syllables, the official orthography does not mark

vowel length and melody. Only the context can tell the reader which word was meant. Written texts are thus ambiguous even to native speakers. Thus, the written word *gusura* can stand either for (*gusura*) 'to fart' or (*gusuura*) 'to visit', and *gutaka* can stand for either (*gutaka*) 'to scream' or (*gutaaka*) 'to decorate'; *ino* can stand for either (*ino*) 'toe' or (*inó*) 'here', *inda* can stand for (*inda*) 'stomach' or (*indá*) 'louse', *umuryango* can stand for (*umuryaango*) 'family' or (*umuryáango*) 'door', and *ikirere* can stand for (*ikireere*) 'banana leaf' or (*ikiréeré*) 'air space'. Even though the sound 'p' has been lost and is found only in onomatopoeic words and loan words, the aspirated voiceless velar fricative 'h' is spelled as 'p' after the bilabial nasal 'm', as shown in the examples *impuha* (*imhuuha*) 'rumors', *impamvu* (*imhaámvu*) 'cause/reason'. The allophones, the voiced bilabial stop 'b', which appears only after the homorganic nasal 'm', and the voiced bilabial fricative a, realized intervocalically, are also written the same way, using the voiced bilabial stop symbol 'b'. Although the language has only one liquid, both 'r' and 'l' are used in the orthography. The liquid 'r' is used in all texts and 'l' is used only in loan words that have 'l' in their spelling, such as *Libiya* 'Libya', *Alijeriya* 'Algeria', *dolari* 'dollar'.

Vowels and Consonants

Kinyarwanda has five vowels, which are either long or short and are high-toned or have no tones. The high tone can appear on either the first mora or the second mora. These vowels are the two high vowels 'i' and 'u', the midvowels 'e' and 'o', and the center low vowel 'a'. The midvowels 'e' and 'o' are not allowed in both the (pre)prefix and suffix positions. In verbs, however, these midvowels can appear in the suffix position as a result of vowel harmony if the vowel of the verb stem is a midvowel, e.g., *gukosa* 'to make mistakes', *gukosoora* 'to correct' /ku-kos-uur-a/; *kumenya* 'to know', *kumenyeesha* /ku-meny-iish-a/.

The majority of word stems have the same identical vowel in all syllables: *u-mu-biri* 'body', *u-bu-riri* 'bed, *i-ki-reenge* 'leg', *i-béere* 'breast', *u-mu-góongo* 'back', *u-mu-hoondo* 'yellow', *u-ku-guru* 'leg', *u-ru-túgu* 'shoulder', *igihaánga* 'skull', *i-ki-gaanza* 'hand palm'. This observation raises the question as to whether the stem is assigned only one vowel that is copied or that spreads to other syllables.

Since Kinyarwanda has open syllables only, loan words with cluster consonants copy the vowel of the syllable on the right or a default vowel 'u' with bilabial consonants and 'i' with other consonants. As the following loan word *porogaramu* 'program' shows, both vowels 'o' and 'a' are copied on the preceding vowelless consonants and the vowel 'u' is inserted after the final consonant 'm'.

'program' /p.$ro.$g.$ra.$m.$/ > *porogaramu*

This language has both simple and complex consonants. The simple consonants, using the official orthography, are the bilabials 'p', 'b', and 'm', the interdentals 'f' and 'v', the alveolars 't', 'd', 's', 'z', and 'n'; the alveopalatals 'sh', and 'j', and the velars 'k', 'g', and 'h'. Kinyarwanda has two glides, the palatal 'y' and the bilabial 'w'. It has one liquid, 'r', which is written as 'l' in some loan words, as was pointed out earlier. The affricates are the labiodental 'pf', the alveolar 'ts', and the palatal 'c'.

The complex consonants are the prenasalized simple consonants, the palatalized consonants, the velarized consonants, the palatalized–velarized consonants, the prenasalized palatalized consonants, and the prenasalized palatalized–velarized consonants.

Prenasalized consonants are the bilabial 'mp' and 'mb'; the interdental 'mv', 'mf', and 'mpf'; the alveolar 'nt', 'nd', 'ns', 'nz', and 'nts'; the palatal 'nsh', 'nj', and 'nc'; and the velar 'nk', 'ng', and 'nshy'.

Palatalized consonants are the bilabial 'by', 'py', and 'my'; the interdental 'fy'; the alveolar 'ty', 'dy', 'sy', and 'nny'; and the velar 'cy', 'jy', and 'shy'.

The velarized consonants are the bilabial 'pw', 'bw', and 'mw'; the interdental 'fw' and 'vw'; the alveolar 'tw', 'dw', 'sw', 'zw', 'nw', 'rw', and 'tsw'; the palatal 'shw', 'jw', 'cw', and 'yw'; and the velar 'kw', 'gw', and 'hw'. Palatalized–velarized consonants are the bilabial 'byw', 'pyw', and 'myw'; the alveolars 'tyw', 'dyw', 'syw'; and the velar fricative 'shyw'. Palatalized consonants, velarized consonants, and palatalized–velarized consonants can in turn be prenasalized as shown in the following examples: 'mbyw' (prenasalized palatalized–velarized voiced bilabial), 'mvyw' (prenasalized palatalized–velarized voiced interdental, 'nshyw' prenasalized palatalized–velarized voiceless fricative velar, and 'njyw' (prenasalized palatalized–velarized voiced stop velar). The complex consonants in Kinyarwanda are discussed at great length in Kimenyi (2002) and Bizimana *et al.* (1998). It is still an open debate in phonetics and phonology as to whether these complex consonants are one with multiple articulators or a sequence of independent segments.

Tonology

Role of Tones

Tones are lexical, morphological, and syntactical. Lexical tones differentiate words that look alike segmentally as shown in (1), morphological tones play

the role that segmental morphemes are assigned in other languages as illustrated in (2), whereas syntactic tones are assigned depending on where the word bearing the tone occurs in the noun phrase, verb phrase or the sentence as shown in (3):

(1) *inda* 'stomach' <> *indá* 'louse', *ino* 'toe' <> *inó* 'here'

(2) *basoma* 'they read' <> *basomá* 'who read' <> *básoma* 'when they read'

(3) *baraaza bagakóra* 'they come and work' <> *baraaza bagakora akazi* 'they come and do the work'

In (2), the lack of tone shows the present tense, and the high tone on the second syllable shows that the verb is a relative clause, whereas the high tone on the first syllable of the verb stem shows that the verb is a temporal or conditional clause. In (3), the verb loses its high tone because it is followed by a complement.

Tone Rules

Tone rules in Kinyarwanda were thought to be complicated. However, when looked at very closely, they are very simple. There is only one lexical high tone per morpheme. Some morphemes are toneless. Noun tone patterns differ from verb tone rules. Any noun can have a lexical high tone on any syllable of the stem, except the augment and the prefix. A verb, however, even when it is polysyllabic and has multiple suffixes, can have a high tone only on the first mora of the first syllable or the first of the second syllable of the stem. Other syllables are extraprosodic. When high tones are found there, they are stray tones, since they do not participate in tone rules such as the Meeussen rule, Beat Movement, Iambic Reversal, etc. The first syllable verb high tone assignment is lexical, whereas the second syllable high tone assignment is grammatical. The prosodic domain of tone rules application of both nouns and verbs is the leftmost phonological tone and the first mora of the stem first syllable. Tone rules apply from right to left, whereas in the majority of languages whose tone rules have been studied, they apply from left to right.

Nouns can obtain a secondary and a tertiary tone. A secondary high tone is assigned on the first mora of the noun stem if the lexical high tone is at least two or more mora away from the first noun mora.

isáandukú 'box' ← /i-saandukú/, *inkókorá* 'elbow' ← /i-n-kokorá/
aug-box aug-CL9-elbow
abasásamígozí 'murderers'
/a-ba-sas-a+i-mi-gozí/
aug-sub.pr.-make bed-asp+aug-CL4-rope
inshóberamáhaánga 'idiomatic expressions'
/i-n-shober-a+a-ma-haánga/
aug-CL9-disorient-asp+aug-CL6-foreign countries

A noun can thus have only a maximum of three phonetic high tones.

What makes verb tones seem complex is the assignment of the tense–aspect–modality morphemes. Some tenses or moods erase lexical tones, thus making the whole finite verb toneless, or assign tones to toneless verb stems, making both toneless verb stems and high-toned verb stems look the same. As shown below, the verb stem *-kin-* 'play' and *-kór-* 'work/do' are neutralized, becoming toneless or both bearing a high tone in some tenses.

ntibagikora 'they do not work anymore' <> *ntibagikina* 'they do not play anymore'
baracyáakóra 'they still work' <> *baracyáakina* 'they still play'
bakoré 'they should work'<> *bakiné* 'they should play'.

The metrical domain for verb tone rules is the first mora of the first object pronoun and the first mora of the verb stem for lexical tones and the first mora of the first object pronoun and the first mora of the second syllable of the verb stem. Kinyarwanda is one of the Bantu languages that can have multiple object pronouns.

baranáhabíbamákoreeshereza ← /ba-ra-na-ha-bi-ba-mu–kór-iish-ir-ir-y-a/
they-t-also-there-it-him/her-them-do-appl-caus-appl-appl-caus-asp
'they also make them do it for him/her there'.

Phonology

Phonological Rules Affecting Vowels

Phonological rules affecting vowels are vowel deletion, vowel coalescence, gliding, vowel harmony, vowel shortening, and vowel lengthening.

When a word or morpheme that ends with a vowel is followed by another one that also starts with a vowel, the final vowel of the word or morpheme on the left is always deleted.

Vowel coalescence takes place within a word if there is a sequence of two morphemes ending with the central low vowel 'a' and starting with the high vowels 'i' and 'u', respectively: 'a + i' becomes 'e' and 'a + u' becomes 'o', e.g., *améenyo* /a-ma-íinyo/ 'teeth'.

Gliding takes place within words or clitics if in a sequence of two vowels the first one is a high vowel ('i', 'u') or a round vowel ('u', 'o'), thus becoming 'y' for the front high vowel and 'w' for round vowels, e.g., *harimó amáazi* > harim$^{w'}$áamáazi 'there is water

in it'; *i-ki-úuma* > ik^{y}úuma 'knife'; /*u-bu-oónko*/ > ub^{w}oónko 'brain'. Vowel harmony affects vowels in the suffix position. If the suffix vowel is high ('i', 'u') it becomes mid ('e', 'o', respectively), if the word stem vowel is a mid vowel, e.g., *gukóra* /ku-kór-a/ 'to work' > *gukóreesha* /ku-kór-iish-a/ 'to cause to work/ employ/use', *kumenya* /ku-meny-a/ 'to know' > *kumenyeesha*/ku-meny-iish-a/ 'to let know/inform'. Vowels in Kinyarwanda are always short in the beginning and final positions of words. They always lengthen before prenasalized consonants and after palatalized and labio-velarized consonants (Kimenyi, 1979, 2002).

Phonological Rules Affecting Consonants

Phonological rules affecting consonants are assimilation, dissimilation, known as Dahl's Law, fricative spread, deletion, and insertion. As in other Bantu languages, reduplication is also very productive in Kinyarwanda.

Consecutive consonants acquire the same voice, manner, and place of articulation phonetic features. Nasals take the place of articulation (labial, velar, palatal, velar) of the consonant on the right. Consonants obtained through the palatalization or velarization process also agree in voice, nasality, and place of articulation with the governor consonant.

gukúbita 'to hit' > *gukúbitkwa* 'to be hit' /ku-kubit-w-a/

kudóda 'to sew' > *kudódgwa* 'to be sewed' /ku-dód-w-a/

kubóna 'to see' > *kubónnwa* 'to be seen' /ku-bón-w-a/.

If a word has a palatalized fricative in one of the syllables on the right, fricatives in preceding syllables become palatalized as well.

gusoonza /ku-soonz-a/ 'to be hungry' > *gushoonjeesha* /ku-soonz-iish-a/ 'to cause hunger'

basuuzugura /ba-suuzugur-a/ 'they despise' > *bashuujuguje* /ba-suuzugur-ye/ 'they just caused to despise'.

This phenomenon argues for the autosegmental treatment of phonological rules because as the provided examples show, these fricatives do not have to be in adjacent syllables.

Reduplication is both lexical and grammatical. Lexical reduplication consists of stems that are already reduplicated. Grammatical reduplication affects the stem. In verbs, it is very productive with verbs of movement or sound to show repetition, iterativity, or intensity.

Reduplication is achieved by either repeating the fist syllable or the whole stem.

gutuumba 'to swell' → *gututuumba* 'to start swelling'

kugeenda 'to go'/'walk' → *kugeendageenda* 'to walk around'

ukwéezi 'moon/month' → *icyéezeezí* 'moonlight'

ubusá 'nothing' → *ubusáabusá* 'very little quantity'.

Morphology

Noun Morphology

Kinyarwanda has 16 classes. Modifiers (adjectives, demonstratives, numerals, possessives) agree with the head noun by taking this class marker.

In some cases, however, the class marker has different phonetic forms depending on the grammatical category of the modifier, as illustrated in Table 1. The sentence in (2) in Table 1 shows how this type of noun class agreement works. The head noun *abagabo* (a-ba-gabo) 'men' with class 2 prefix-ba-has it copied to all modifying elements (adjectives, subject pronouns, object pronouns, etc.).

bá-no	*ba-gabo*	*ba-tatu*	*ba-gufi*,
these	men	three	short

mu-ra-bá-bon-a,	ba-mez-e nééza b-óose.
you-pres-them-see-asp	they-are-asp well all

'These three short men, you see them, they are all of them doing well".

Table 1 Allomorphic variation of the nominal prefix according to its function

	Noun	*Adjective*	*Object pronoun*	*Demonstive*	*Possessive*
1.	u-mu-	mu-	-mu-	u-	u-
2.	a-ba-	ba-	-ba-	ba-	ba-
3.	u-mu-	mu-	-wu-	u-	u-
4.	i-mi-	mi-	-yi-	i-	i-
5.	i-ri-	ri-	-ri-	ri-	ri-
6.	a-ma-	ma-	-ya-	a-	a-
7.	i-ki-	ki-	-ki-	ki-	ki-
8.	i-bi	bi-	-bi-	bi-	bi-
9.	i-n-	n-	-yi-	i-	i-
10.	i-n-	n-	-zi-	zi-	zi-
11.	u-ru-	ru-	-ru-	ru-	ru-
12.	a-ka-	ka-	-ka-	ka-	ka-
13.	u-tu-	tu-	-tu-	tu-	tu-
14.	u-bu-	bu-	-bu-	bu-	bu-
15.	u-ku-	ku-	-ku-	ku-	ku-
16.	a-ha-	ha-	-ha-	ha-	ha-

Note: The numbers 1–16 correspond to traditional conventional Bantu noun classification.

The Use of the Preprefix The preprefix or augment usually does not have any semantic function. Some Bantu languages such as Kiswahili do not have it. In Kinyarwanda, it is deleted, after demonstratives, in the vocative case and in onomastics (name creation). Within certain words, however, its absence marks definiteness and its presence indefiniteness.

mugaanga 'the doctor' > *umugaanga* 'a doctor'
munywáanyi 'the buddy' > *umunywáanyi* 'a buddy'
mugeenzi 'the friend' > *mugeenzi* 'a friend'
mwaarimú 'the teacher' > *umwáarimú* 'a teacher'

The absence of the preprefix bleeds tone rules and its absence feeds them. The secondary tone assignment on the first more of the noun stem takes place only if it has a preprefix.

Noun Derivation with the Prefix Kinyarwanda nouns have a small number of suffixes. The most productive one is *-kazi*, which is added to the stem to show feminine, e.g., *umunyarwaanda* 'Rwandan'<> *umunyarwaandakazi* 'female Rwandan', *umwáarimú* 'teacher' <> *umwáarimúkazi* 'female teacher'.

Derivation is productive with the preprefix that creates new words that are either metaphorically or metonymically related to the original noun as shown by the stem-ntu in the following examples: *umuuntu* 'person', *ikiintu* 'object', *ukuuntu* 'manner', *ubuuntu* 'generosity', *ahaantu* 'place'.

Verb Morphology

The simple Kinyarwanda verb form consists of the subject pronoun, the verb stem, and the aspect marker. The aspect marker is either -a(ga) (imperfective aspect) or -ye (the perfective aspect as seen below. The -aga suffix is used in past tenses only and is not used in Kirundi.

basoma (ba-som-a) 'they read';
basomaga /ba-a-som-aga/ 'they were reading';
basomye (ba-som-ye) 'they just read'.

The complex form consists of the preprefix, the subject pronoun, the tense–aspect–modality (TAM) morphemes, the object markers, the reflexive pronoun -i-, the verb stem, the lexical verb extensions, the grammatical morphemes, the aspect marker and the postsuffixes -mó, -hó, or -yó.

The preprefixes are either the morpheme nti- or ni-, negative and temporal morphemes, respectively. The TAM morphemes show time, mood or aspect. Two TAM morphemes can occur in the same slot. Kinyarwanda can have multiple object pronouns, multiple lexical verbal extensions, and multiple grammatical suffixes. Lexical extensions such as -agur-, -iir-, uur, -aang, iriz-, etc., add lexical information, such as inchoativity, iterativity, repetitivity, intensity, frequentativity, reversivity. Grammatical morphemes, such as the causative morpheme -iish-, the applicative morpheme -ir-, the comitative/reciprocal morpheme -an-, can be added to any verb stem. The following sentence serves as an example to illustrate a verb with multiple object pronouns and multiple grammatical suffixes:

Umugoré a-ra-na1-ha2-ki3-zi4-ba5-ku6-n7- som-eesh-eesh-er-er-eza
woman she-pres-also1-there2-it3-it4-them5-you6-me7-read-caus-caus-appl-appl-asp
'The woman is also making them use it to do it for me for me there'.

Lack of Adjectives Kinyarwanda has a handful of adjectives (less than 20). What is expressed by adjectives in other languages is rendered by either the possessive construction (X of Y) or the relative clause construction.

'a poor person' :	*umuuntu*	*w'úmukené*
	person	of poor
	umuuntu	*ukénnye*
	person	who-is poor

Ideophones Ideophones are common not only in Bantu languages but also in the whole Niger–Congo language family. They are different from onomatopoeias, which imitate sounds of nature. They can express different concepts that do not have anything to do with sound by using sound symbolism, short or long vowels, reduplication, triplication, or quadruplication. They can also have different grammatical functions.

umuseké *weerá de*.
dawn which-is clear (ideophone)
'a very clear dawn'
icyáayi *tsiritsiri*
tea ideophone
'a very dark coffee'.

Unclassified Categories As was pointed out earlier, Kinyarwanda has a handful of adjectives. It is the same with function words as well, namely, auxiliaries, prepositions, conjunctions, and subordinators. These are expressed by noun phrases or verb phrases. In most cases, the structure tells whether the noun or verb is the noun or the verb or a function word.

muu nsií y'{á}am{é}ez{á}
under earth of table
'under the table'

baravúgana usíibye kó batabonána.
they-talk-to-each-other you-are-absent that they-do not see each other
'they talk to each other except that they do not see each other'.

Stems of nouns, verbs, and unclassified words can have different phonetic variations, as the word for diploma shows:

diploma: *dipóroómi*, *dipóromá*, *dipóromé*, *dipóromí*, *dipóromó*, *dipóromú*, *dipóroóma*, *dipóroóme*, *dipóroómu*, *dipóroómo*, *diipóroómu*, *diipóroómo*, *diipóroóme*, *diipóroómi*, *diipóroóma*.

It is still an unsolved question for Kinyarwanda lexicographers to decide which form should be considered the main form or whether all forms should entered in the dictionary as independent lexical entries.

Syntax

Kinyarwanda, like other Bantu languages, is a SVO language. Modifiers follow head nouns. What is interesting about this language, as pointed out in Kimenyi (1980, 2002), is the existence of (a) the subject–object reversal, (b) the wh-question *in situ*, (c) the lack of relative pronouns, (d) serialization, and (e) the existence of multiple direct objects.

Object–Subject Reversal and Existential Construction

The object–subject reversal consists of interchanging the object and the subject positions, whereas the existential construction puts both the subject and the object after the verb, prefixing the verb with the locative morpheme ha-(CL16). Neither construction changes the meaning, except that focus is on the object.

Umwáana a-ra-som-a igitabo.
'The child is reading the book'.
child s/he-t-read-asp book
Igitabo ki-ra-som-a umwáana.
'The book is reading the child'.
book CL7-t-read-asp child
Ha-ra-som-a igitabo umwáana.
'It is the child who is reading the book'.
CL16-t-read-asp book child

The object–subject reversal and the existential constructions have the same function as the passive, which is shown by the suffix -w- added to the verb just before the aspect marker.

Igitabo ki-ra-som-w-a n'úumwáana.
'The book is being read by the child.'
book CL7-t-read-pass-asp by child

Wh-Question *In Situ*

In Kinyarwanda and many other Bantu languages, wh-questioning is only allowed *in situ*.

W-iit-w-a nde?
you-call-pass-asp who
'you are called who?' > 'What is your name?'

Ba-tuu-ye he?
they-live-asp where
'Where do they live?'

Lack of Relative Pronouns

Kinyarwanda does not have relative pronouns. Relative constructions are marked by a high tone on the verb stem instead.

Abáana ba-som-á ibitabo.
children they-read-asp/rel books
'The children who read books'.
Ibitabo abáana ba-som-á.
books children they-read-asp/rel
'The books that the children read'.

Serial Verb Construction

When multiple verbs precede the sentence main verb, they lose their semantic function and serve as auxiliaries or tense–aspect–modality bearers. This is illustrated by the following sentences:

Ba-a-ri	*bá-tuu-ye*	*bá-saanz-w-e*
they-t-be	they-dwell-asp	they-join-pass-asp
aux^1	aux^2	aux^3
bá-jy-a	*bá-kuund-a*	*gu-pf-a*
they-go-asp	they-like-asp	to-die
aux^4	aux^5	aux^6

ku-dú hamagar-a.
to-us-call-asp
V
'They usually at least called us'.

Mu-siga-ye	*mú-geend-a*	*mú-heerako*
you-stay-asp	you-walk-asp	you-start
aux^1	aux^2	aux^3

mú-du-subiz-a.
from you-us-answer-asp
V
'Now you respond to us immediately'.

Multiple Direct Objects

Kinyarwanda, like many other Bantu languages, can have multiple direct objects. These objects are either inherent or structural.

Recipients or benefactives are introduced directly to the verb without any preposition with some interactive verbs (giving, showing, etc.).

Umugabo a-haa-ye abáana ibiryó
man he-give-asp children food
'The man has just given food to the children'.

Umwáarimú a-r-éerek-a abanyéeshuúri amashusho.
teacher he-t-show-asp students pictures
'The teacher is showing pictures to the students'.

Inalienable possessions can appear as direct objects without any verb extension marker.

Umugoré a-ra-kúbit-a umwáana ukuguru n'ínkoni.
woman sub.pr.-t-hit-asp child leg with stick
Umugoré a-ra-kúbit-a umwáana inkoni ku kuguru.
woman sub.pr.-t-hit-asp child stick on leg
'The woman hitting the child on the leg with a stick'.

As shown by these examples, the inalienable possession '*ukuguru*' and the instrumental '*inkoni*' can appear as either adjuncts or direct objects without any verbal extension.

Structural direct objects are obtained by deleting prepositions of adjunct objects and by adding suffixes such as -iish-, -ir-, or -an- to the verb stem.

Umugabo a-ra-andik-a ibáruwá n'íikáramú
man sub.pr.-t-write-asp letter with pen
Umugabo a-ra-andik-iish-a ikáramú ibáruwá
man sub.pr.-t-write-caus-asp pen letter
'The man is writing a letter with a pen'.

Conclusion

Kinyarwanda is a prototypical Bantu language. It has all the features that characterize this language group. Its main contributions in syntax have been about the nature and function of grammatical relations (Kimenyi, 1980) and in tonology, its contributions have been about the nature of tone, tone representations, and tone rule application (Kimenyi, 2002).

See also: Bantu Languages; Burundi: Language Situation; Congo, Democratic Republic of: Language Situation; Ideophones; Meeussen, Achille Emile (1912–1978); Phonetics of Harmony Systems; Rwanda: Language Situation; Serial Verb Constructions; Tanzania: Language Situation; Uganda: Language Situation.

Bibliography

Bizimana S *et al.* (1998). *Imiteérere y'Ikinyarwaanda.* Kigali, Rwanda: Pelloti Set Press.

Coupez A (1980). *Abrégé de Grammaire rwanda. Tome I and Tome II.* Butare, Rwanda: Institut National de Recherch Scientifique.

Kimenyi A (1979). *Studies in Kinyarwanda and Bantu Phonology.* Edmonton, Alberta: Linguistic Research Inc.

Kimenyi A (1980). *A relational grammar of Kinyarwanda.* Berkeley/Los Angeles: University of California Press.

Kimenyi A (2002). *A tonal grammar of Kinyarwanda: An autosegmental and metrical analysis.* Lewiston, NY: The Edwin Mellen Press.

Kirghiz

L Johanson, Mainz University, Mainz, Germany

Location and Speakers

Kirghiz (**qïrγïz tili, qïrγïzča**) belongs to the Northwestern or Kipchak branch of the Turkic language family, more specifically, to its Southern or Aralo-Caspian group. Until the early 20th century, it was called Kara-Kirghiz, whereas Kazakh was referred to as Kirghiz or Kazak-Kirghiz. Kirghiz is spoken in the Kyrgyz Republic (**Qïrγïz Respublikasï**) or Kyrgyzstan (**Qïrγïzstan**) and in parts of Uzbekistan, Tajikistan, China (Xinjiang), the Russian Federation, Kazakhstan, etc. Its main area is the mountainous part of Western Turkistan, the plateaus of the western Tienshan south of Kazakhstan, and the Alay mountain south of Ferghana. The number of speakers amounts to about 3 million, in Kyrgyzstan over 2.5 million.

In spite of the existence of a modern Kirghiz standard language, Russian has remained the dominant language of higher education, administration, and so forth in the Republic. Since Kirghiz was proclaimed the official language of Kyrgyzstan in 1989, it has consolidated its position, acquiring more social functions. In 1996, Russian was made an official language, along with Kirghiz, in territories and workplaces in which Russian-speaking citizens predominate.

Origin and History

It is still unclear to what extent the Kirghiz of today are successors of the Old Kirghiz, the first Turkic people mentioned in Chinese sources and described

Umugabo a-haa-ye abáana ibiryó
man he-give-asp children food
'The man has just given food to the children'.

Umwáarimú a-r-éerek-a abanyéeshuúri amashusho.
teacher he-t-show-asp students pictures
'The teacher is showing pictures to the students'.

Inalienable possessions can appear as direct objects without any verb extension marker.

Umugoré a-ra-kúbit-a umwáana ukuguru n'ínkoni.
woman sub.pr.-t-hit-asp child leg with stick
Umugoré a-ra-kúbit-a umwáana inkoni ku kuguru.
woman sub.pr.-t-hit-asp child stick on leg
'The woman hitting the child on the leg with a stick'.

As shown by these examples, the inalienable possession '*ukuguru*' and the instrumental '*inkoni*' can appear as either adjuncts or direct objects without any verbal extension.

Structural direct objects are obtained by deleting prepositions of adjunct objects and by adding suffixes such as -iish-, -ir-, or -an- to the verb stem.

Umugabo a-ra-andik-a ibáruwá n'íikáramú
man sub.pr.-t-write-asp letter with pen
Umugabo a-ra-andik-iish-a ikáramú ibáruwá
man sub.pr.-t-write-caus-asp pen letter
'The man is writing a letter with a pen'.

Conclusion

Kinyarwanda is a prototypical Bantu language. It has all the features that characterize this language group. Its main contributions in syntax have been about the nature and function of grammatical relations (Kimenyi, 1980) and in tonology, its contributions have been about the nature of tone, tone representations, and tone rule application (Kimenyi, 2002).

See also: Bantu Languages; Burundi: Language Situation; Congo, Democratic Republic of: Language Situation; Ideophones; Meeussen, Achille Emile (1912–1978); Phonetics of Harmony Systems; Rwanda: Language Situation; Serial Verb Constructions; Tanzania: Language Situation; Uganda: Language Situation.

Bibliography

Bizimana S *et al.* (1998). *Imiteérere y'Ikinyarwaanda.* Kigali, Rwanda: Pelloti Set Press.

Coupez A (1980). *Abrégé de Grammaire rwanda. Tome I and Tome II.* Butare, Rwanda: Institut National de Recherch Scientifique.

Kimenyi A (1979). *Studies in Kinyarwanda and Bantu Phonology.* Edmonton, Alberta: Linguistic Research Inc.

Kimenyi A (1980). *A relational grammar of Kinyarwanda.* Berkeley/Los Angeles: University of California Press.

Kimenyi A (2002). *A tonal grammar of Kinyarwanda: An autosegmental and metrical analysis.* Lewiston, NY: The Edwin Mellen Press.

Kirghiz

L Johanson, Mainz University, Mainz, Germany

Location and Speakers

Kirghiz (**qïrγïz tili, qïrγïzča**) belongs to the Northwestern or Kipchak branch of the Turkic language family, more specifically, to its Southern or Aralo-Caspian group. Until the early 20th century, it was called Kara-Kirghiz, whereas Kazakh was referred to as Kirghiz or Kazak-Kirghiz. Kirghiz is spoken in the Kyrgyz Republic (**Qïrγïz Respublikasï**) or Kyrgyzstan (**Qïrγïzstan**) and in parts of Uzbekistan, Tajikistan, China (Xinjiang), the Russian Federation, Kazakhstan, etc. Its main area is the mountainous part of Western Turkistan, the plateaus of the western Tienshan south of Kazakhstan, and the Alay mountain south of Ferghana. The number of speakers amounts to about 3 million, in Kyrgyzstan over 2.5 million.

In spite of the existence of a modern Kirghiz standard language, Russian has remained the dominant language of higher education, administration, and so forth in the Republic. Since Kirghiz was proclaimed the official language of Kyrgyzstan in 1989, it has consolidated its position, acquiring more social functions. In 1996, Russian was made an official language, along with Kirghiz, in territories and workplaces in which Russian-speaking citizens predominate.

Origin and History

It is still unclear to what extent the Kirghiz of today are successors of the Old Kirghiz, the first Turkic people mentioned in Chinese sources and described

there as blond and blue-eyed. This group settled on Upper Yenisey. Runiform inscriptions found on the territory of today's Tuva indicate that the first Kirghiz state was, at the beginning of the 8th century A.D., located north of the Sayan mountains. In 840 the Kirghiz ended the old steppe Uyghur empire and established their own empire, which lasted until 920. Most old Turkic groups left this region at the turn of the millennium. A few groups remained in Siberia, (e.g., the ancestors of the Kirghiz and the Altay Turks).

Some Kirghiz tribes may already have moved to the Tienshan region by the 10th century. Other tribes followed, particularly during the Mongol attacks. The Mongol expansion in the 13th century forced old Kirghiz groups to migrate to Western Turkistan and the Tienshan region.

In the following centuries, the tribes referred to as Kirghiz were gradually pushed back by Oirats and Dzungars. In the 16th century, the Kirghiz acted as an important ally of the Kazakh. In the second half of the 17th century, the Yenisey Kirghiz were forced to accept the sovereignty of the Kalmyk. At the beginning of the 18th century, the majority of the Kirghiz migrated to Tienshan. After the breakdown of the Dzungar Empire in 1758, the Kirghiz definitively settled in their present territory. In the 18th and 19th centuries, they were subjects of the Uzbek Khanate of Kokand. They came under Russian supremacy in 1867, and their territory was incorporated into the Russian Empire in 1880. Under the Russian and Soviet rule, numerous Kirghiz emigrated to China. In 1991, the Kirghiz Republic was proclaimed an independent state.

Related Languages and Language Contacts

Kirghiz is closely related to Southern Altay Turkic of South Siberia. Its modern form is also very close to Kazakh as a result of long-standing intensive contacts. Old Kirghiz, as attested to in inscriptions, was similar to Orkhon Turkic and Old Uyghur (*see* **Turkic Languages**). The language was later influenced by Mongolic and especially by Kipchak Turkic. In the 18th and 19th centuries, it was subject to some impact from the Iranicized dialects of the Uzbek area. The contacts with Russian began at the end of the 19th century. After the Kirghiz territory was conquered by the Russian empire in the second half of the 19th century, the Russian influence became predominant.

The Written Language

In the Soviet period, a Kirghiz standard language was developed on the basis of the northern dialects. Before the revolution Kirghiz had already found some limited use as a written language. A modified version of the Arabic script was introduced in 1924 but was given up in 1928 in favor of the unified Roman-based alphabet. A modified Cyrillic-based script was introduced in 1940. In the post-Soviet era, a new Roman-based alphabet was created, but it has not yet replaced the Cyrillic-based script. The Arabic script is used for the variety written in China.

Distinctive Features

Kirghiz exhibits most linguistic features typical of the Turkic family (*see* **Turkic Languages**). It is an agglutinative language with suffixing morphology, sound harmony, and a head-final constituent order. In the following, only a few distinctive features will be dealt with. In the notation of suffixes, capital letters indicate phonetic variation; for example, **A** = **a/e,** and **I** = **ï/i.** Hyphens are used here to indicate morpheme boundaries.

Phonological Features

Kirghiz has a rather regular vowel system, lacking the reduced vowels of Kazakh. The different orthographic vowel representations very often conceal the phonetic similarities between Kirghiz and Kazakh.

Kirghiz exhibits the typical Kipchak labialization of **g/γ**, but does not, as Kazakh, replace them by the glide **w**. As in Altay Turkic, the preceding vowel is deleted and lengthened, as in **to:-lu:** [mountain-DER] 'mountainous' (< **ta:γ-l°γ**), cf. Altay Turkic **tu:-lu:** [mountain-DER], Kazakh **taw-lï** [mountain-DER]. Kirghiz differs from Kazakh by absence of the sound changes **č** > **š** and **š** > **s**. Modern Kirghiz displays initial **ǰ**- instead of older **y**- (e.g., **ǰol** 'way'; cf. Turkish **yol**). This change occurred under the influence of Kazakh, which, however, later changed **ǰ to ž**. Kirghiz does not display the uvular **x** or the glottal **h** but replaces them by **q** or zero in loanwords (e.g., **qabar** 'message' [< **xaber**], **ar** 'each' [< **har**]). Kirghiz has, like some Siberian Turkic languages, a well-developed sound harmony. Suffixes exhibit both front vs. back harmony and rounded vs. unrounded harmony. The choice of suffix vowels is determined by features of the preceding syllable (e.g., **köl-dör-dön** [lake-PL-ABL] 'from the lakes,' **üy-lör-übüz-dö** [house-PL-POSS.1.PL-LOC] 'in our houses'). The rounded vs. unrounded harmony also affects low suffix vowels, which are rounded to **o** and **ö** after **o**, **ö**, **ü** in the preceding syllable. However, **u** in the preceding syllable is not followed by **o** (e.g., **qum-da** [sand-LOC] 'in the sand'). Rounding of low suffix vowels is also

observed in Altay Turkic, Turkmen, Bashkir, and to some extent, in Kazakh, and Noghay.

Morphophonemic alternations are found in suffixes with initial **l** and **n**, which are assimilated to **d** after voiced consonants (with some exceptions after **r** and **y**), and to **t** after voiceless consonants (e.g., **ata-nïn** [father-GEN] 'of the father,' **qar-dïn** [snow-GEN] 'of the snow,' **at-tïn** [horse-GEN] 'of the horse,' **alma-lar** [apple-PL] 'apples,' **kün-dör** [day-PL] 'days,' **at-tar** [horse-PL] 'horses').

Arabic and Persian loanwords have generally been adapted more strongly to the native phonological system than their counterparts in neighboring Turkic languages (e.g., **apta** 'week' [< **hafta**], **ubaqtï** 'time' [< **waqt**]).

Grammar

The Kirghiz genitive suffix -**nIn** ends in -**n**, as in languages of the Southwestern (Oghuz) branch, not in -**ŋ** as in the neighboring languages.

The suffix -**rA:K** (after consonant-final stems -**IrA:K**) is added to adjectives to express comparative degree. The superlative is expressed by means of the preposed element **eŋ** (e.g., **eŋ ǰaqšï** [SUPERL good] 'best'). The normal second-person plural form of the personal pronoun is **si-ler** [you-PL] 'you.' **Siz** is the polite form used for one addressee, and **siz-der** [you-PL] is the polite form for more than one addressee. The corresponding copula suffixes are -**slŋAr**, -**slz** and -**slz-dAr**. Most demonstrative pronouns exhibit an optional -**l** in the nominative; for example, **bu(l)** 'this.' Different degrees of closeness are marked with **bu(l)** (close referents known to the speaker), **ošo(l)** and **ušul** (more remote referents), and **tigi(l)** and **teti-gil(l)** (remote referents outside the conversational setting).

Pronouns also include the reflexive **öz** and the interrogatives **emne** 'what,' **kim** 'who,' **qaysï** 'which,' and so on. A collective suffix -**O:** (with drop of stem-final vowels) is used with the numerals from 1 to 7 (e.g., **ek-ö:** [two-COLL] 'two together'). Distributives are formed with the ablative (e.g., **otuz-dan** [thirty-ABL] 'thirty each'). Approximative and multiplicative numbers can be expressed with -**dAy**, -**čA** and -**LA-GAn** (e.g., **otuz-day** [thirty-APPR], **otuz-ča** [thirty-APPR] 'about thirty,' **ǰüz-dö-gön** [hundred-MULT] 'hundreds').

As in most other Turkic languages, the third-person possessive suffix exhibits the so-called 'pronominal **n**' (e.g., dative -**n-A**, locative -**n-dA**). The past copula particle is **ele** 'was,' instead of **edi** ~ **idi** in other Turkic languages.

The cooperative-reciprocal form in -**(I)š** is also used to indicate the third-person plural of finite verbs (e.g., **ǰaz-ïš-at** [write-PL-PAST.3] 'they write' [singular **ǰaz-at**]), **kel-iš-ti** [come-PL-PAST.3] 'they came' [singular **kel-di** [come-PAST.3.SG]). A general present tense is formed with -**A** + personal markers (e.g., **bar-a-t** [go-PRES-3.SG] 'goes'). A more focal present tense (with a narrower focus on the ongoing event) is formed with- convert + present forms of **ǰat**- 'to lie' (and three other auxiliary verbs) + personal markers (e.g., **bar-a ǰat-a-t** [go-CONV AUX-PRES-3.SG] 'is going'). The suffix -**Dlr** adds a presumptive meaning (e.g., **oyγon-γon-dur** [wake up-POSTTERMINAL-PRES.3.SG] 'has presumably waken up'). An evidential past is formed with- **(I)p-tlr**. A habitual or durative past is formed with- **čl** + personal markers (e.g., **kel-čü-büz** [come-HABIT.PAST-1.PL] 'we used to come'). Intention is expressed by -**MAK-čI** (e.g., **kel-mek-či-min** [come-INTENT-1.SG] 'I want to come'). The system of conjunctions is weakly developed, as Kirghiz has not been under strong Iranian influence.

Lexicon

The basis of the Kirghiz vocabulary consists of Kipchak Turkic elements. Similar to the languages of other nomadic groups, Kirghiz has borrowed numerous Mongolic words as a result of close contacts, especially in the Middle Ages (e.g., **dülöy**, 'deaf'; **belen**, 'ready'; **qara-**, 'to look'). Arabic and Persian words, copied via Chaghatay and Uzbek particularly into the southern dialects, constitute a sizable part of the vocabulary, covering various domains of Islamic culture (e.g., **künö**, 'sin'; **ša:r**, 'city'; **baqča**, 'garden'; **pikir**, 'thought'). The northern dialects, on which the literary language is based, are less influenced by the Islamic vocabulary.

Russian loanwords, which constitute the most recent layer in the lexicon, were introduced from the end of the 19th century on. The use of Russian words is rather dominant in informal spoken standard Kirghiz. Efforts have been made in the last decades to reduce the amount of Russian terms by creating neologisms on the basis of Turkic and Arabic–Persian lexical material or on Russian models. Neologistic suffixes include -**čIl** (e.g., **ulut-čul** [nation-DER] 'nationalist'; from **ulut**, 'nation'), -**GIč** (e.g., **uč-quč** [fly-DER] 'pilot' [from **uč**- 'to fly']).

The variety of Kirghiz spoken in Xinjiang exhibits numerous Chinese loanwords. The written language, however, is largely oriented toward the norm used in Kirghizstan.

Dialects

The Kirghiz dialects can be divided into a southern and a northern group, the latter forming the base

of the standard language. Northern dialects tend toward intervocalic voicening of **s** (e.g., **bala-zï** [child-POSS.3.SG] 'his/her child' instead of **bala-sï** [child-POSS.3.SG]), a feature typical of the South Siberian Turkic languages. The southern dialects, mainly those spoken in the Ferghana basin, show different degrees of Uzbek influence. They lack some characteristic features that separate Kirghiz from other Turkic languages. For example, the glide **w** is preserved in cases in which the standard language only exhibits a long vowel as a trace of a velar that once underwent labialization (e.g., **tow** 'mountain' instead of **to:**) The past copula particle **ede** 'was' is used instead of **ele**. The plural suffix is used in third-person plural forms of finite verb (e.g., **kel-di-ler** [come-PAST-PL.3] 'they came'), instead of **kel-iš-ti** [come-PL-PAST.3]. The 'pronominal **n**' has been lost under Southeastern Turkinc influence. There are also other changes in the nominal inflection. Words copied from Arabic and Persian via Chaghatay and Uzbek have preserved their original phonetic shape to a higher degree than in the northern dialects.

See also: Altaic Languages; China: Language Situation; Kazakh; Kazakhstan: Language Situation; Kyrgyzstan: Language Situation; Russian Federation: Language Situation; Tajikistan: Language Situation; Turkic Languages; Uzbekistan: Language Situation.

Bibliography

Hebert R J & Poppe N (1963). *Kirghiz manual.* The Hague: Mouton.

Imart G (1981). Le kirghiz*: 1-2.* Aix-en-Provence: Publications de l'Université de Provence.

Johanson L (2001). 'Kirghiz.' In Garry J & Rubino C (eds.) *Facts about the world's major languages: an encyclopedia of the world's major languages, past and present.* New York: H. W. Wilson, New England Publishing Associates. 387–389.

Judaxin K K (1965). *Kirgizsko-russkij slovar'.* Moskva: Sovetskaja Ėnciklopedija.

Kirchner M (1998). 'Kirghiz.' In Johanson L & Csató É Á (eds.) *The Turkic languages.* London: Routledge. 344–356.

Wurm S (1949). 'The (Kara-)Kirghiz language.' *Bulletin of the School of Oriental Studies 13,* 97–120.

Zaxarova O V (ed.) (1987). *Grammatika kirgizskogo literaturnogo jazyka* 1: *Fonetika i morfologija*. Frunze: Ilim.

Kiribati: Language Situation

B Palmer, University of Surrey, Guildford, UK

The Republic of Kiribati is a chain of 33 atolls and tiny coral islands straddling the equator in Micronesia in the central-western Pacific north of Fiji. The total land area is about 811 square kilometers scattered across over 5 million square kilometers of ocean.

Kiribati was initially settled over 2000 years ago by Austronesian speakers. The Kiribati remained skilled ocean voyagers, and there is evidence of regular contact with other Pacific groups. However, Kiribati remained isolated from Western influence until the arrival of whalers and missionaries during the 19th century. Formally taken under British control as the Gilbert Islands in 1877, Kiribati became independent in 1979. About one-third of its population lives in the national capital, Bairiki, on Tarawa atoll.

The country is home to one indigenous language spoken by over 97% of its approximately 100 000 inhabitants. The Kiribati language, also known as Gilbertese, is a Nuclear Micronesian language within the Oceanic branch of Austronesian. It is closely related to Ponapean, Kosraean, and Marshallese. The linguistic origin of the Micronesian grouping is not certain, but it is likely to have been in the Solomon Islands or Vanuatu. Kiribati is the most conservative Nuclear Micronesian language, which suggests (along with geography) that it was settled very early in the Austronesian spread into Micronesia.

Kiribati has two main dialects. One is associated with islands in the north of the group, and the other with those in the south. No standard variety has been adopted, but in practice the southern dialect is emerging as a standard. An additional dialect is spoken on Makin and Butaritari, and another on Nui (in Tuvalu). The entire speech community of a third dialect, Banaban, was relocated from Banaba (Ocean Island) to Rabi Island in Fiji in the 1950s. Smaller relocated iKiribati communities are found in Nauru, the Solomon Islands, Vanuatu, and elsewhere in Fiji.

The Kiribati language retains a central role in modern Kiribati. It predominates in newspapers, on the radio, and in the churches. The language also has official recognition, but takes second place to English. The national language is Kiribati, while English is the official language. The country's constitution, laws, and official government documents have both a Kiribati and an English text, but the English is

of the standard language. Northern dialects tend toward intervocalic voicening of **s** (e.g., **bala-zï** [child-POSS.3.SG] 'his/her child' instead of **bala-sï** [child-POSS.3.SG]), a feature typical of the South Siberian Turkic languages. The southern dialects, mainly those spoken in the Ferghana basin, show different degrees of Uzbek influence. They lack some characteristic features that separate Kirghiz from other Turkic languages. For example, the glide **w** is preserved in cases in which the standard language only exhibits a long vowel as a trace of a velar that once underwent labialization (e.g., **tow** 'mountain' instead of **to:**) The past copula particle **ede** 'was' is used instead of **ele**. The plural suffix is used in third-person plural forms of finite verb (e.g., **kel-di-ler** [come-PAST-PL.3] 'they came'), instead of **kel-iš-ti** [come-PL-PAST.3]. The 'pronominal **n**' has been lost under Southeastern Turkinc influence. There are also other changes in the nominal inflection. Words copied from Arabic and Persian via Chaghatay and Uzbek have preserved their original phonetic shape to a higher degree than in the northern dialects.

See also: Altaic Languages; China: Language Situation; Kazakh; Kazakhstan: Language Situation; Kyrgyzstan: Language Situation; Russian Federation: Language Situation; Tajikistan: Language Situation; Turkic Languages; Uzbekistan: Language Situation.

Bibliography

Hebert R J & Poppe N (1963). *Kirghiz manual.* The Hague: Mouton.

Imart G (1981). Le kirghiz*: 1-2.* Aix-en-Provence: Publications de l'Université de Provence.

Johanson L (2001). 'Kirghiz.' In Garry J & Rubino C (eds.) *Facts about the world's major languages: an encyclopedia of the world's major languages, past and present.* New York: H. W. Wilson, New England Publishing Associates. 387–389.

Judaxin K K (1965). *Kirgizsko-russkij slovar'.* Moskva: Sovetskaja Ėnciklopedija.

Kirchner M (1998). 'Kirghiz.' In Johanson L & Csató É Á (eds.) *The Turkic languages.* London: Routledge. 344–356.

Wurm S (1949). 'The (Kara-)Kirghiz language.' *Bulletin of the School of Oriental Studies 13,* 97–120.

Zaxarova O V (ed.) (1987). *Grammatika kirgizskogo literaturnogo jazyka* 1: *Fonetika i morfologija*. Frunze: Ilim.

Kiribati: Language Situation

B Palmer, University of Surrey, Guildford, UK

The Republic of Kiribati is a chain of 33 atolls and tiny coral islands straddling the equator in Micronesia in the central-western Pacific north of Fiji. The total land area is about 811 square kilometers scattered across over 5 million square kilometers of ocean.

Kiribati was initially settled over 2000 years ago by Austronesian speakers. The Kiribati remained skilled ocean voyagers, and there is evidence of regular contact with other Pacific groups. However, Kiribati remained isolated from Western influence until the arrival of whalers and missionaries during the 19th century. Formally taken under British control as the Gilbert Islands in 1877, Kiribati became independent in 1979. About one-third of its population lives in the national capital, Bairiki, on Tarawa atoll.

The country is home to one indigenous language spoken by over 97% of its approximately 100 000 inhabitants. The Kiribati language, also known as Gilbertese, is a Nuclear Micronesian language within the Oceanic branch of Austronesian. It is closely related to Ponapean, Kosraean, and Marshallese. The linguistic origin of the Micronesian grouping is not certain, but it is likely to have been in the Solomon Islands or Vanuatu. Kiribati is the most conservative Nuclear Micronesian language, which suggests (along with geography) that it was settled very early in the Austronesian spread into Micronesia.

Kiribati has two main dialects. One is associated with islands in the north of the group, and the other with those in the south. No standard variety has been adopted, but in practice the southern dialect is emerging as a standard. An additional dialect is spoken on Makin and Butaritari, and another on Nui (in Tuvalu). The entire speech community of a third dialect, Banaban, was relocated from Banaba (Ocean Island) to Rabi Island in Fiji in the 1950s. Smaller relocated iKiribati communities are found in Nauru, the Solomon Islands, Vanuatu, and elsewhere in Fiji.

The Kiribati language retains a central role in modern Kiribati. It predominates in newspapers, on the radio, and in the churches. The language also has official recognition, but takes second place to English. The national language is Kiribati, while English is the official language. The country's constitution, laws, and official government documents have both a Kiribati and an English text, but the English is

defined as the authoritative version. Parliament may be conducted in Kiribati or English, although in practice Kiribati is normally used, mixed with specialist English terminology. Kiribati is used exclusively in education in the earliest years, giving way to English at the secondary level. A government Kiribati Language Board was established in the 1970s to standardize orthography and grammar, prepare a dictionary (yet to appear), and encourage Kiribati literature, although vernacular literature remains limited.

See also: Austronesian Languages: Overview; Language Education Policies in the Pacific; Malayo-Polynesian Languages.

Bibliography

Groves T, Groves G & Jacobs R (1985). *Kiribatese: an outline description.* Canberra: Pacific Linguistics.

Sabatier E (1971). *Gilbertese-English dictionary.* Oliva M (trans.). Sydney: South Pacific Commission.

Klatt, Dennis H. (1938–1988)

K N Stevens

This article is reproduced from the previous edition, volume 4, p. 1852, © 1994, Elsevier Ltd.

Born on March 31, 1938, Dennis Klatt carried out his undergraduate and early graduate studies in electrical engineering at Purdue University. In 1964, he completed his doctorate in the communication sciences program at the University of Michigan, with a thesis entitled 'Theories of aural physiology.' He spent most of his professional career at the Massachusetts Institute of Technology, where he was for many years a senior research scientist in the Department of Electrical Engineering and Computer Science, and in the Research Laboratory of Electronics. His activities at MIT and his role as a consultant to industry consisted of a balanced combination of fundamental research into the mechanisms of human speech production and perception and applied research leading to practical devices. He died on December 30, 1988.

Klatt's fundamental contributions to an understanding of speech processes covered a broad range of topics. His detailed experimental studies of speech timing in English led to several papers on the rules governing control of the durations of the sounds, words, and phrases that make up spoken sentences. He had a continuing interest in the problem of how speech is processed in the auditory system and central nervous system of human listeners. This speech perception research ranged from detailed experimental studies of how listeners extract information from the rapidly changing events in speech to models of the storage and access of lexical information by listeners. In his publications, he argued convincingly that the development of appropriate distance metrics is a central problem in determining whether two utterances represent two tokens of the same word or represent two different words. His contributions to the advancement of automatic speech recognition included serving on the advisory committee of a national effort in this area in the 1970s and publication of a comprehensive review of the field as it existed at that time.

The area where Klatt's work had a major impact for those working in the field of speech communication is speech synthesis. He developed a flexible computer-based synthesizer that can be used to generate sounds in basic studies of human speech perception. The software for this synthesizer was made available to research laboratories that are active in these areas, making possible new quantitative research on the processes of auditory and speech perception. The versatility of the speech synthesis system lies in its ability to generate in a natural way a variety of utterances with different amounts of nasalization, with aspiration, with different glottal characteristics, and so on. With this synthesizer as a base, Klatt developed a complete system for synthesis of speech from English orthography. This system, implemented as DECtalk, was generally acknowledged to produce the most intelligible speech of any synthesizer available at the time. He carried out a detailed study of the characteristics of the voice source, particularly for female talkers, in an effort to improve further the naturalness of the synthesizer output.

Bibliography

Klatt D H (1976). 'The linguistic uses of segmental duration in English: Acoustic and perceptual evidence.' *Journal of the Acoustical Society of America 59,* 1208–1221.

defined as the authoritative version. Parliament may be conducted in Kiribati or English, although in practice Kiribati is normally used, mixed with specialist English terminology. Kiribati is used exclusively in education in the earliest years, giving way to English at the secondary level. A government Kiribati Language Board was established in the 1970s to standardize orthography and grammar, prepare a dictionary (yet to appear), and encourage Kiribati literature, although vernacular literature remains limited.

See also: Austronesian Languages: Overview; Language Education Policies in the Pacific; Malayo-Polynesian Languages.

Bibliography

Groves T, Groves G & Jacobs R (1985). *Kiribatese: an outline description.* Canberra: Pacific Linguistics.

Sabatier E (1971). *Gilbertese-English dictionary.* Oliva M (trans.). Sydney: South Pacific Commission.

Klatt, Dennis H. (1938–1988)

K N Stevens

This article is reproduced from the previous edition, volume 4, p. 1852, © 1994, Elsevier Ltd.

Born on March 31, 1938, Dennis Klatt carried out his undergraduate and early graduate studies in electrical engineering at Purdue University. In 1964, he completed his doctorate in the communication sciences program at the University of Michigan, with a thesis entitled 'Theories of aural physiology.' He spent most of his professional career at the Massachusetts Institute of Technology, where he was for many years a senior research scientist in the Department of Electrical Engineering and Computer Science, and in the Research Laboratory of Electronics. His activities at MIT and his role as a consultant to industry consisted of a balanced combination of fundamental research into the mechanisms of human speech production and perception and applied research leading to practical devices. He died on December 30, 1988.

Klatt's fundamental contributions to an understanding of speech processes covered a broad range of topics. His detailed experimental studies of speech timing in English led to several papers on the rules governing control of the durations of the sounds, words, and phrases that make up spoken sentences. He had a continuing interest in the problem of how speech is processed in the auditory system and central nervous system of human listeners. This speech perception research ranged from detailed experimental studies of how listeners extract information from the rapidly changing events in speech to models of the storage and access of lexical information by listeners. In his publications, he argued convincingly that the development of appropriate distance metrics is a central problem in determining whether two utterances represent two tokens of the same word or represent two different words. His contributions to the advancement of automatic speech recognition included serving on the advisory committee of a national effort in this area in the 1970s and publication of a comprehensive review of the field as it existed at that time.

The area where Klatt's work had a major impact for those working in the field of speech communication is speech synthesis. He developed a flexible computer-based synthesizer that can be used to generate sounds in basic studies of human speech perception. The software for this synthesizer was made available to research laboratories that are active in these areas, making possible new quantitative research on the processes of auditory and speech perception. The versatility of the speech synthesis system lies in its ability to generate in a natural way a variety of utterances with different amounts of nasalization, with aspiration, with different glottal characteristics, and so on. With this synthesizer as a base, Klatt developed a complete system for synthesis of speech from English orthography. This system, implemented as DECtalk, was generally acknowledged to produce the most intelligible speech of any synthesizer available at the time. He carried out a detailed study of the characteristics of the voice source, particularly for female talkers, in an effort to improve further the naturalness of the synthesizer output.

Bibliography

Klatt D H (1976). 'The linguistic uses of segmental duration in English: Acoustic and perceptual evidence.' *Journal of the Acoustical Society of America 59,* 1208–1221.

Klatt D H (1980). 'Software for a cascade/parallel formant synthesizer.' *Journal of the Acoustical Society of America 67,* 971–995.

Klatt D H (1987). 'Review of text-to-speech conversion for English.' *Journal of the Acoustical Society of America 82,* 737–793.

Klatt D H & Klatt L C (1990). 'Voice quality variation within and across female and male talkers: Implications for speech analysis, synthesis, and perception.' *Journal of the Acoustical Society of America 87,* 820–857.

Knorosow, Yuri (1922–1999)

O Molchanova, Uniwersytet Szczecinski, Szczecin, Poland

Yuri Knorosow, a Russian ethnographer of Central America and linguist, was born on November 19, 1922 in Kharkov (the Ukrainian republic), and died in 1990 in Moscow. He graduated from Moscow University in 1948 and obtained his Ph.D. in 1955. Knorosow's most substantial contribution to scholarship is the decipherment of Mayan hieroglyphic writings, and also ancient Indian inscriptions from Mohenjo-Daro (an archeological site in Pakistan, near the Indus River) and Harappa (the site of successive cities of the Indus Valley civilization in Pakistan). Research on these Indian inscriptions revealed their correlation in grammatical structure with the Dravidian family.

Mayan inscriptions date back to the 1st century A.D., and the Mayan writing system employs 300 signs. There are four Mayan codices left in the world, and linguists have worked out a grid of syllabic glyphs in Maya as well as a significant portion of the logograms. The Mayan people wrote both logographically and phonetically, and logograms will often have phonetic complements to facilitate their correct identification.

The first attempts to decipher Maya were in the middle of the 19th century. In the middle of the 1950–1960s, Knorosow contributed greatly by making use of positional statistics, which opened up the possibility of studying Maya diachronically.

See also: Dravidian Languages; Mayan Languages.

Bibliography

Knorozov J V (1955a). 'Vvodnaja statja, primečanija i perevod.' In Landa D (ed.) *Soobščenija o delakh v Jukatane.* Moskva-Leningrad: Nauka.

Knorozov J V (1955b). *Sistema pis'ma drevnikh maja.* Moskva: Nauka.

Knorozov J V (1963). *Pis'mennost' indejcev maja.* Moskva-Leningrad: Nauka.

Knorozov J V (1975). *Ieroglifičeskije rukopisi maja.* Leningrad: Nauka.

Koasati *See:* Muskogean Languages.

Kobayashi, Hideo (1903–1978)

R Washio, University of Tsukuba, Ibaraki, Japan

Japanese linguist Hideo Kobayashi, best known as the translator of *Cours de linguistique générale*, was born in Tokyo on January 5, 1903. He graduated from Tokyo Imperial University in 1927 with the thesis *Le langage de Henrik Ibsen: Essai stylistique*. When he did the translation of the *Cours*, he was only 24 years old and was already described as a rising scholar and a natural-born linguist who had mastered more than 10 modern European languages, "not to speak of Classical Greek, Latin and Sanskrit" (Kindaichi, 1928: 521).

Klatt D H (1980). 'Software for a cascade/parallel formant synthesizer.' *Journal of the Acoustical Society of America 67,* 971–995.
Klatt D H (1987). 'Review of text-to-speech conversion for English.' *Journal of the Acoustical Society of America 82,* 737–793.
Klatt D H & Klatt L C (1990). 'Voice quality variation within and across female and male talkers: Implications for speech analysis, synthesis, and perception.' *Journal of the Acoustical Society of America 87,* 820–857.

Knorosow, Yuri (1922–1999)

O Molchanova, Uniwersytet Szczecinski, Szczecin, Poland

Yuri Knorosow, a Russian ethnographer of Central America and linguist, was born on November 19, 1922 in Kharkov (the Ukrainian republic), and died in 1990 in Moscow. He graduated from Moscow University in 1948 and obtained his Ph.D. in 1955. Knorosow's most substantial contribution to scholarship is the decipherment of Mayan hieroglyphic writings, and also ancient Indian inscriptions from Mohenjo-Daro (an archeological site in Pakistan, near the Indus River) and Harappa (the site of successive cities of the Indus Valley civilization in Pakistan). Research on these Indian inscriptions revealed their correlation in grammatical structure with the Dravidian family.

Mayan inscriptions date back to the 1st century A.D., and the Mayan writing system employs 300 signs. There are four Mayan codices left in the world, and linguists have worked out a grid of syllabic glyphs in Maya as well as a significant portion of the logograms. The Mayan people wrote both logographically and phonetically, and logograms will often have phonetic complements to facilitate their correct identification.

The first attempts to decipher Maya were in the middle of the 19th century. In the middle of the 1950–1960s, Knorosow contributed greatly by making use of positional statistics, which opened up the possibility of studying Maya diachronically.

See also: Dravidian Languages; Mayan Languages.

Bibliography

Knorozov J V (1955a). 'Vvodnaja statja, primečanija i perevod.' In Landa D (ed.) *Soobščenija o delakh v Jukatane.* Moskva-Leningrad: Nauka.
Knorozov J V (1955b). *Sistema pis'ma drevnikh maja.* Moskva: Nauka.
Knorozov J V (1963). *Pis'mennost' indejcev maja.* Moskva-Leningrad: Nauka.
Knorozov J V (1975). *Ieroglifičeskije rukopisi maja.* Leningrad: Nauka.

Koasati *See:* Muskogean Languages.

Kobayashi, Hideo (1903–1978)

R Washio, University of Tsukuba, Ibaraki, Japan

Japanese linguist Hideo Kobayashi, best known as the translator of *Cours de linguistique générale*, was born in Tokyo on January 5, 1903. He graduated from Tokyo Imperial University in 1927 with the thesis *Le langage de Henrik Ibsen: Essai stylistique*. When he did the translation of the *Cours*, he was only 24 years old and was already described as a rising scholar and a natural-born linguist who had mastered more than 10 modern European languages, "not to speak of Classical Greek, Latin and Sanskrit" (Kindaichi, 1928: 521).

Klatt D H (1980). 'Software for a cascade/parallel formant synthesizer.' *Journal of the Acoustical Society of America 67,* 971–995.
Klatt D H (1987). 'Review of text-to-speech conversion for English.' *Journal of the Acoustical Society of America 82,* 737–793.
Klatt D H & Klatt L C (1990). 'Voice quality variation within and across female and male talkers: Implications for speech analysis, synthesis, and perception.' *Journal of the Acoustical Society of America 87,* 820–857.

Knorosow, Yuri (1922–1999)

O Molchanova, Uniwersytet Szczecinski, Szczecin, Poland

Yuri Knorosow, a Russian ethnographer of Central America and linguist, was born on November 19, 1922 in Kharkov (the Ukrainian republic), and died in 1990 in Moscow. He graduated from Moscow University in 1948 and obtained his Ph.D. in 1955. Knorosow's most substantial contribution to scholarship is the decipherment of Mayan hieroglyphic writings, and also ancient Indian inscriptions from Mohenjo-Daro (an archeological site in Pakistan, near the Indus River) and Harappa (the site of successive cities of the Indus Valley civilization in Pakistan). Research on these Indian inscriptions revealed their correlation in grammatical structure with the Dravidian family.

Mayan inscriptions date back to the 1st century A.D., and the Mayan writing system employs 300 signs. There are four Mayan codices left in the world, and linguists have worked out a grid of syllabic glyphs in Maya as well as a significant portion of the logograms. The Mayan people wrote both logographically and phonetically, and logograms will often have phonetic complements to facilitate their correct identification.

The first attempts to decipher Maya were in the middle of the 19th century. In the middle of the 1950–1960s, Knorosow contributed greatly by making use of positional statistics, which opened up the possibility of studying Maya diachronically.

See also: Dravidian Languages; Mayan Languages.

Bibliography

Knorozov J V (1955a). 'Vvodnaja statja, primečanija i perevod.' In Landa D (ed.) *Soobščenija o delakh v Jukatane.* Moskva-Leningrad: Nauka.
Knorozov J V (1955b). *Sistema pis'ma drevnikh maja.* Moskva: Nauka.
Knorozov J V (1963). *Pis'mennost' indejcev maja.* Moskva-Leningrad: Nauka.
Knorozov J V (1975). *Ieroglifičeskije rukopisi maja.* Leningrad: Nauka.

Koasati *See:* Muskogean Languages.

Kobayashi, Hideo (1903–1978)

R Washio, University of Tsukuba, Ibaraki, Japan

Japanese linguist Hideo Kobayashi, best known as the translator of *Cours de linguistique générale*, was born in Tokyo on January 5, 1903. He graduated from Tokyo Imperial University in 1927 with the thesis *Le langage de Henrik Ibsen: Essai stylistique*. When he did the translation of the *Cours*, he was only 24 years old and was already described as a rising scholar and a natural-born linguist who had mastered more than 10 modern European languages, "not to speak of Classical Greek, Latin and Sanskrit" (Kindaichi, 1928: 521).

In 1929, Kobayashi was appointed to the Faculty of Law and Literature at Keijo Imperial University, which was established in Seoul, Korea by Imperial Japan in 1924. With the end of World War II, he returned to Japan. In 1949, he was appointed Professor at Tokyo Institute of Technology. After he retired from Tokyo Tech in 1961, he served as Professor at Waseda University until his retirement in 1973.

Kobayashi's *Gengogaku genron* (Fundamentals of Linguistics) was the first translation of the *Cours*, published three years before the German translation appeared in 1931. Its title was changed to *Ippan gengogaku kōgi* (Course in General Linguistics) when it underwent the last major revision in 1972—five years before Kobayashi died. His translation, which may still be the only translation of the *Cours* in a non-Indo-European language, made the *Cours* accessible to all those capable of reading Japanese, including a large number of Korean scholars and students. The words Kobayashi coined for notions such as *synchronie* and *diachronie* have become firmly established in both Japan and Korea.

Saussure's influence in Japan has been described as "immense" (*Édition critique*, p. 369). De Mauro suggested that a factor contributing to it might have been the prior influence in Japan of E. D. Polivanov and H. E. Palmer. It was unquestionably Kobayashi, however, who had played a decisive role in the spread and penetration of the Saussurean thoughts in Japan. As early as the 1930s, more than 20 years prior to the publication of the English translation of the *Cours*, Saussure's theory of language had been met in Japan with a harsh criticism from Motoki Tokieda, a scholar in the traditional school of Japanese linguistics. Like many of his contemporaries, Tokieda had learned about Saussure's theory through Kobayashi's translation of the *Cours*, and even more importantly, by Kobayashi's own writings on linguistics, some of which had served as a kind of companion to (the Japanese version of) the *Cours* in the early days of modern linguistics in Japan. In addition to the *Cours*, Kobayashi had also translated such works as Louis Hjelmslev's *Principes de grammaire générale* and Charles Bally's *Linguistique générale et linguistique française*.

See also: De Mauro, Tullio (b. 1932); Japan: History of Linguistics; Kindaichi, Kyosuke (1882–1971); Palmer, Harold Edward (1877–1949); Saussure, Ferdinand (-Mongin) de (1857–1913); Tokieda, Motoki (1900–1967).

Bibliography

Kindaichi K (1928). '*Gengogaku genron*-o yomu [Review of Kobayashi (1928)].' *Minzoku 3,* 519–521.

Kobayashi H (1928). *Gengogaku genron* [Japanese translation of Saussure (1916)]. Tokyo: Oka shoin. Rev. version published as *Ippan gengogaku kōgi.* Tokyo: Iwanami shoten, 1972.

Saussure F de (1916). *Cours de linguistique générale. Édition critique* by T. de Mauro. Paris: Payot, 1975.

Kodzu, Harushige (1908–1973)

T Maruyama, Nanzan University, Nagoya, Japan

Harushige Kodzu was a pioneer scholar of Indo-European comparative linguistics and ancient Greek, especially its Arcadian dialect, in 20th century Japan.

Having graduated from the University of Tokyo in 1930, Kodzu left Japan for the United Kingdom to study comparative grammar from 1931 to 1934 at Oxford University under the guidance of various professors, including Frederick W. Thomas and E. Braunholtz. Upon returning to Japan, he continued his research in the graduate school of the University of Tokyo. In 1936 he married Kumiko Honma and had a daughter. The contributions of his wife and his daughter as philological assistants to his research were considerable. In 1938 he was appointed research assistant of the department of linguistics at the University of Tokyo, and lecturer in 1942. He was granted his D.Litt. in 1946 from the same university. Two years later, he was appointed associate professor of linguistics, and he was promoted to full professor in 1951. In March 1968, he retired from the university with the title of Emeritus Professor. In the following year, he was invited to become the Dean of the Faculty of Letters of Musashi University. He continued his work until his death in 1973. His major publications are listed in *Gengo Kenkyu* (the Journal of the Linguistic Society of Japan) No. 64, published in 1973, with an obituary written by his disciple Kiyozo Kazama.

In his doctoral dissertation of the ancient Greek Arcadian dialect (1946), he meticulously studied carved inscriptions and linguistic boundaries with other dialects. He carried out extensive and thorough investigations of stone inscriptions within the linguistic

In 1929, Kobayashi was appointed to the Faculty of Law and Literature at Keijo Imperial University, which was established in Seoul, Korea by Imperial Japan in 1924. With the end of World War II, he returned to Japan. In 1949, he was appointed Professor at Tokyo Institute of Technology. After he retired from Tokyo Tech in 1961, he served as Professor at Waseda University until his retirement in 1973.

Kobayashi's *Gengogaku genron* (Fundamentals of Linguistics) was the first translation of the *Cours*, published three years before the German translation appeared in 1931. Its title was changed to *Ippan gengogaku kōgi* (Course in General Linguistics) when it underwent the last major revision in 1972—five years before Kobayashi died. His translation, which may still be the only translation of the *Cours* in a non-Indo-European language, made the *Cours* accessible to all those capable of reading Japanese, including a large number of Korean scholars and students. The words Kobayashi coined for notions such as *synchronie* and *diachronie* have become firmly established in both Japan and Korea.

Saussure's influence in Japan has been described as "immense" (*Édition critique*, p. 369). De Mauro suggested that a factor contributing to it might have been the prior influence in Japan of E. D. Polivanov and H. E. Palmer. It was unquestionably Kobayashi, however, who had played a decisive role in the spread and penetration of the Saussurean thoughts in Japan. As early as the 1930s, more than 20 years prior to the publication of the English translation of the *Cours*, Saussure's theory of language had been met in Japan with a harsh criticism from Motoki Tokieda, a scholar in the traditional school of Japanese linguistics. Like many of his contemporaries, Tokieda had learned about Saussure's theory through Kobayashi's translation of the *Cours*, and even more importantly, by Kobayashi's own writings on linguistics, some of which had served as a kind of companion to (the Japanese version of) the *Cours* in the early days of modern linguistics in Japan. In addition to the *Cours*, Kobayashi had also translated such works as Louis Hjelmslev's *Principes de grammaire générale* and Charles Bally's *Linguistique générale et linguistique française*.

See also: De Mauro, Tullio (b. 1932); Japan: History of Linguistics; Kindaichi, Kyosuke (1882–1971); Palmer, Harold Edward (1877–1949); Saussure, Ferdinand (-Mongin) de (1857–1913); Tokieda, Motoki (1900–1967).

Bibliography

Kindaichi K (1928). '*Gengogaku genron*-o yomu [Review of Kobayashi (1928)].' *Minzoku 3,* 519–521.

Kobayashi H (1928). *Gengogaku genron* [Japanese translation of Saussure (1916)]. Tokyo: Oka shoin. Rev. version published as *Ippan gengogaku kōgi.* Tokyo: Iwanami shoten, 1972.

Saussure F de (1916). *Cours de linguistique générale. Édition critique* by T. de Mauro. Paris: Payot, 1975.

Kodzu, Harushige (1908–1973)

T Maruyama, Nanzan University, Nagoya, Japan

Harushige Kodzu was a pioneer scholar of Indo-European comparative linguistics and ancient Greek, especially its Arcadian dialect, in 20th century Japan.

Having graduated from the University of Tokyo in 1930, Kodzu left Japan for the United Kingdom to study comparative grammar from 1931 to 1934 at Oxford University under the guidance of various professors, including Frederick W. Thomas and E. Braunholtz. Upon returning to Japan, he continued his research in the graduate school of the University of Tokyo. In 1936 he married Kumiko Honma and had a daughter. The contributions of his wife and his daughter as philological assistants to his research were considerable. In 1938 he was appointed research assistant of the department of linguistics at the University of Tokyo, and lecturer in 1942. He was granted his D.Litt. in 1946 from the same university. Two years later, he was appointed associate professor of linguistics, and he was promoted to full professor in 1951. In March 1968, he retired from the university with the title of Emeritus Professor. In the following year, he was invited to become the Dean of the Faculty of Letters of Musashi University. He continued his work until his death in 1973. His major publications are listed in *Gengo Kenkyu* (the Journal of the Linguistic Society of Japan) No. 64, published in 1973, with an obituary written by his disciple Kiyozo Kazama.

In his doctoral dissertation of the ancient Greek Arcadian dialect (1946), he meticulously studied carved inscriptions and linguistic boundaries with other dialects. He carried out extensive and thorough investigations of stone inscriptions within the linguistic

island of the Arcadian dialect, and based on his extremely detailed research concerning the boundaries with other dialects, he came to entertain misgivings about the then generally accepted theory of P. Kretchmer that the Ionians were the oldest Greek ancestors. He arrived at the conclusion that the Ionians evolved from the mixture of several races of Asia Minor.

Kodzu's attitude toward his research was extremely catholic, in that he did not confine himself only to the dimension of language; rather, his vision embraced all aspects of the classical world, which he sought to understand as a whole. In his approach to the reconstruction of prehistoric Greek languages, he took into account various legends, archeological discoveries, as well as place names, to obtain a truly comprehensive understanding.

Kodzu was a leading committee member of the Linguistic Society of Japan ever since its foundation in 1938. His contribution to classical philology in Japan is also remarkable. He worked for *Seiyo Koten Kenkyu* (Journal of Classical Studies) as one of the editors from its foundation in 1953 until his death. By translating a number of Greek classical literary works into Japanese and publishing more than 20 books concerning Indo-European linguistics or the Greek language, he planted the genuine tradition of western classical philology in Japan.

See also: Greek, Ancient; Japan: History of Linguistics; Kazama, Kiyozo (b. 1928); Latin.

Bibliography

Kazama K (1973). 'Harushige Kodzu 1908–73.' *Gengo Kenkyu 64,* 3–11.

Kodzu H (1937). *Greek studies, vol. I. On the dialect of Alkman; vol. II, A study of instrumental adverbs, particles, and conjunctions, those in -α-,-η-,-ω-.* Tokyo: Iwanami, Shoten.

Kodzu H (1950). 'Origin of the Ionians.' *Gengo Kenkyu 15,* 64–79.

Kodzu H (1951). 'Absorption of Arcadian into κοινή.' *Gengo Kenkyu 17/18,* 108–124.

Kodzu H (1954). *Studies in Arcadian dialect* (written in Japanese). Tokyo: Iwanami Shoten.

Kodzu H (1954). *Comparative grammar of Indo-European languages* (written in Japanese). Tokyo: Iwanami Shoten.

Kodzu H (1960). *Greek grammar* (written in Japanese). Tokyo: Iwanami Shoten.

Kodzu H (1960). 'Linguistics in Japan 1945–58.' *Kratylos 5,* 1–9.

Koelle, Sigismund Wilhelm (1823–1902)

S Pugach, Ohio State University, Lima, OH, USA

Sigismund Wilhelm Koelle was born in Kleebronn, Württemberg, on July 14, 1823, and died on February 2, 1902, in London. Koelle was one of a handful of German missionaries from Württemburg, including J. L. Krapf and J. F. Schön, who fundamentally shaped the nascent discipline of African linguistics in the mid-19th century (*see* **Krapf, Johann Ludwig (1810–1881)**). From 1841 to 1845, Koelle studied to become a missionary in Basel, then joined the London-based Church Missionary Society. In 1847, after briefly studying Arabic in Tübingen, Koelle was sent to Freetown, Sierra Leone. While there, Koelle was employed as a linguist at the Fourah Bay Institute. The Institute, later renamed Fourah Bay College, University College of Sierra Leone, was the first place in Africa where African language research was conducted.

It was in Sierra Leone that Koelle undertook his most renowned enterprise, the *Polyglotta Africana*, a classification of over 100 West African languages, culled from a collection of about 200 (Dalby, 1970). Shortly after his arrival, Koelle began to collect data from ex-slaves, as P.E.H. Hair has noted (Hair, 1967). The terms West Atlantic, Mande, Mandingo, and Kru all derive from the *Polyglotta*; indeed, the genetic and typological classification of the Mande language family arose with Koelle's work (Wyrod, 2003). Koelle also identified nine 'South African' language groups that would today be classified as Bantu or semi-Bantu, and among which he recognized a distinct relationship (Dalby, 1970). The *Polyglotta* comprised approximately 300 words and phrases listed in the various languages – which included five dialects of Ibo, among others – and was the most complete comparative work on African languages at that time (Koelle, 1854b; Hair, 1967; Jungraithmayr and Leger, 1983). Hair has observed that Koelle's classification in the *Polyglotta* was very sophisticated, precisely locating most of the languages involved and accurately assessing their relationships to each other (Hair, 1967). The *Polyglotta* is also invaluable to historians, since many of Koelle's informants were ex-slaves and their

island of the Arcadian dialect, and based on his extremely detailed research concerning the boundaries with other dialects, he came to entertain misgivings about the then generally accepted theory of P. Kretchmer that the Ionians were the oldest Greek ancestors. He arrived at the conclusion that the Ionians evolved from the mixture of several races of Asia Minor.

Kodzu's attitude toward his research was extremely catholic, in that he did not confine himself only to the dimension of language; rather, his vision embraced all aspects of the classical world, which he sought to understand as a whole. In his approach to the reconstruction of prehistoric Greek languages, he took into account various legends, archeological discoveries, as well as place names, to obtain a truly comprehensive understanding.

Kodzu was a leading committee member of the Linguistic Society of Japan ever since its foundation in 1938. His contribution to classical philology in Japan is also remarkable. He worked for *Seiyo Koten Kenkyu* (Journal of Classical Studies) as one of the editors from its foundation in 1953 until his death. By translating a number of Greek classical literary works into Japanese and publishing more than 20 books concerning Indo-European linguistics or the Greek language, he planted the genuine tradition of western classical philology in Japan.

See also: Greek, Ancient; Japan: History of Linguistics; Kazama, Kiyozo (b. 1928); Latin.

Bibliography

Kazama K (1973). 'Harushige Kodzu 1908–73.' *Gengo Kenkyu 64,* 3–11.
Kodzu H (1937). *Greek studies, vol. I. On the dialect of Alkman; vol. II, A study of instrumental adverbs, particles, and conjunctions, those in -α-,-η-,-ω-.* Tokyo: Iwanami, Shoten.
Kodzu H (1950). 'Origin of the Ionians.' *Gengo Kenkyu 15,* 64–79.
Kodzu H (1951). 'Absorption of Arcadian into κοινή.' *Gengo Kenkyu 17/18,* 108–124.
Kodzu H (1954). *Studies in Arcadian dialect* (written in Japanese). Tokyo: Iwanami Shoten.
Kodzu H (1954). *Comparative grammar of Indo-European languages* (written in Japanese). Tokyo: Iwanami Shoten.
Kodzu H (1960). *Greek grammar* (written in Japanese). Tokyo: Iwanami Shoten.
Kodzu H (1960). 'Linguistics in Japan 1945–58.' *Kratylos 5,* 1–9.

Koelle, Sigismund Wilhelm (1823–1902)

S Pugach, Ohio State University, Lima, OH, USA

Sigismund Wilhelm Koelle was born in Kleebronn, Württemberg, on July 14, 1823, and died on February 2, 1902, in London. Koelle was one of a handful of German missionaries from Württemburg, including J. L. Krapf and J. F. Schön, who fundamentally shaped the nascent discipline of African linguistics in the mid-19th century (*see* **Krapf, Johann Ludwig (1810–1881)**). From 1841 to 1845, Koelle studied to become a missionary in Basel, then joined the London-based Church Missionary Society. In 1847, after briefly studying Arabic in Tübingen, Koelle was sent to Freetown, Sierra Leone. While there, Koelle was employed as a linguist at the Fourah Bay Institute. The Institute, later renamed Fourah Bay College, University College of Sierra Leone, was the first place in Africa where African language research was conducted.

It was in Sierra Leone that Koelle undertook his most renowned enterprise, the *Polyglotta Africana*, a classification of over 100 West African languages, culled from a collection of about 200 (Dalby, 1970). Shortly after his arrival, Koelle began to collect data from ex-slaves, as P.E.H. Hair has noted (Hair, 1967). The terms West Atlantic, Mande, Mandingo, and Kru all derive from the *Polyglotta*; indeed, the genetic and typological classification of the Mande language family arose with Koelle's work (Wyrod, 2003). Koelle also identified nine 'South African' language groups that would today be classified as Bantu or semi-Bantu, and among which he recognized a distinct relationship (Dalby, 1970). The *Polyglotta* comprised approximately 300 words and phrases listed in the various languages – which included five dialects of Ibo, among others – and was the most complete comparative work on African languages at that time (Koelle, 1854b; Hair, 1967; Jungraithmayr and Leger, 1983). Hair has observed that Koelle's classification in the *Polyglotta* was very sophisticated, precisely locating most of the languages involved and accurately assessing their relationships to each other (Hair, 1967). The *Polyglotta* is also invaluable to historians, since many of Koelle's informants were ex-slaves and their

accounts provide information on which groups were caught up in the slave trade and over which routes slaves were taken (Curtin and Vansina, 1964).

Koelle's brief sub-Saharan African career was prolific. He was only in Sierra Leone for five years, yet in addition to the *Polyglotta* Koelle was able to write a grammar of the Vai language, in which he also described the indigenous Vai alphabet, as well as a grammar of Kanuri (Koelle, 1851, 1854a). Koelle met with Memelu Duwalu Bukele, leader of the Vai and innovator of the Vai script, prior to compiling his grammar. In the grammar, Koelle attempted to compare Vai word roots with the roots from various European languages in order to demonstrate an affinity among their sounds; he did so in part to dispel racist claims that Africans and Europeans did not belong to the same species (Koelle, 1851; Pugach, 2001). Moreover, according to Thomas Geider, Koelle's *African native literature*, a collection of oral folktales in Kanuri, is the first full anthology of African literature ever published (Koelle, 1854a; Geider, 1997). Geider also comments that the folktales Koelle collected were the first African texts that the brothers Grimm, Jacob and Wilhelm, used in their comparative folklore studies (Geider, 1997; Grimm and Grimm, 1856).

After leaving Africa, Koelle concluded all work on sub-Saharan African languages and turned his attention to North Africa and Asia. In 1855 the Church Missionary Society sent him to Damietta, a city in Egypt, and then to Palestine and Turkey. Koelle learned Turkish and conducted research on it, though it is unclear whether he ever published these labors, and he was certainly not as well known in the West Asian language field. Still, in 1856 Koelle received the Volney prize for his overall contributions to linguistics (Jungraithmayr and Leger, 1983).

See also: Bantu Languages; Grimm, Jacob Ludwig Carl (1785–1863); Kanuri; Krapf, Johann Ludwig (1810–1881).

Bibliography

British and Foreign Bible Society Archive, Cambridge University Library, Correspondence of S. W. Koelle.

Curtin P & Vansina J (1964). 'Sources of the nineteenth century slave trade.' *Journal of African History 5(2),* 185–208.

Dalby D (ed.) (1970). *Language and history in Africa: a volume of collected papers presented to the London seminar on language and history in Africa (held at the School of Oriental and African Studies, 1967–69).* New York: Africana.

Geider T (1997). 'The universe of Kanuri oral literature and documentary texts.' In Cyffer N & Geider T (eds.) *Advances in Kanuri scholarship* (Westafrikanische Studien, Band 17). Cologne: Rüdiger Koppe. 157–224.

Grimm J (1856). *Kinder- und Hausmärchen gesammelt durch die Brüder Grimm.* Göttingen: Dietrich.

Hair P E H (1967). *The early study of Nigerian languages: essays and bibliographies.* (West African Language Monograph Series 7.) Cambridge: Cambridge University Press.

Jungraithmayr H & Leger R (1983). 'Koelle, Sigismund Wilhelm.' In Jungraithmayr H & Möhlig W J G (eds.) *Lexikon der Afrikanistik: Afrikanische Sprachen und ihre Erforschung.* Berlin: Dietrich Reimer Verlag. 131–132.

Jungraithmayr H & Mohlig W J G (eds.), *Lexikon der Afrikanistik: Afrikanische Sprachen und ihre Erforschung.* Berlin: Dietrich Reimer Verlag.

Koelle S W (1851). *Outlines of a grammar of the Vei language, together with a Vei-English vocabulary and an account of the discovery and nature of the Vei mode of syllabic writing.* London: Keegan, Paul.

Koelle S W (1854a). *Grammar of the Bórnu or Kanuri language.* London: Church Missionary House.

Koelle S W (1854b). *Polyglotta Africana; or a comparative vocabulary of nearly three hundred words and phrases, in more than one hundred distinct african languages.* London: Church Missionary House.

Miller J (2003). *Missionary zeal and institutional control: organizational contradictions in the Basel mission on the Gold Coast, 1828–1917.* Grand Rapids: William B. Eerdmans.

Pugach S (2001). 'Afrikanistik and colonial knowledge: Carl Meinhof, the missionary impulse, and African language and culture studies in Germany, 1887–1919.' Ph.D. diss., University of Chicago.

Wyrod C T (2003). *'The light on the horizon: N'ko literacy and formal schooling in Guinea.'* Master's thesis, Elliott School of International Affairs, George Washington University, Washington, DC.

Koerner, E. F. Konrad (b. 1939)

J E Joseph, University of Edinburgh, Edinburgh, UK

Ernst Frideryk Konrad Koerner was born in 1939 in Mlewiec near Torun (Thorn), Poland. He studied English and German philology, history of art, pedagogy, philosophy, and applied linguistics at the Universities of Göttingen, Berlin, and Edinburgh, completing a B.Phil. at the Free University of Berlin in 1965 and an M.A. at the Justus Liebig University of Giessen in 1968. From there he went to Simon Fraser University in Vancouver, where in 1971 he completed his Ph.D. thesis, 'Ferdinand de Saussure: Origin and development of his linguistic theory in western studies of language: A critical evaluation of the evolution of Saussurean principles and their relevance to contemporary linguistic theories,' a revised version of which was published in 1973.

The thesis posed a number of audacious challenges to the received history of modern linguistics. The 1960s had witnessed the putting forward by prominent scholars of a series of Saussurean 'precursors,' including the sociologist Durkheim and Gabelentz, a linguist whose work lay outside the Neogrammarian mainstream in which Saussure was trained. Koerner argued that more convincing sources could be found within that mainstream itself: rather than Gabelentz, he pointed to Paul, whose 1880 *Principien der Sprachgeschichte* was regarded by the Neogrammarians as their 'bible,' and Whitney, the American linguist whom they respected and Saussure whom they revered. In the case of Durkheim, Koerner raised strong objections to the fact that his name is never cited anywhere by Saussure, insisting that a methodology that ignores the documentary record can produce 'influences' without any limit or firm basis.

These assertions would persist as themes of Koerner's work in the decades that followed, as would a number of others: the impact on 19th-century linguistics of the natural sciences, the parallels among various 'revolutions' proclaimed in modern linguistics, from Jones's 'discovery' of Indo-European linguistic affinities in 1786 to Bopp's methodological refinements of 1816, the Neogrammarian revolution of 1876, structuralism in the 1920s, and the Chomskyan revolution of the 1960s. Koerner considers the last of these particularly dubious, in view of the demonstrable continuity between Chomsky's early methodology and that of his predecessors Jakobson, Hockett, and Harris.

After completing his doctorate, Koerner occupied posts as research associate at the Universities of Texas and Indiana, then in 1973 returned to Germany to do research and teach at the University of Regensburg. In 1976 he took up the professorship of linguistics at the University of Ottawa, which he held until his retirement in 1999, when his 60th birthday was marked by three *Festschrift* publications.

Koerner's importance within modern linguistics owes as much to his organizational energies and editorial skills as to his own scholarship. In 1973 he launched *Historiographia linguistica*, the first journal devoted to the history of linguistics, and the book series *Studies in the history of the language sciences*, both with John Benjamins of Amsterdam. In 1978, he organized the first International Conference on the History of the Language Sciences at the University of Ottawa; this has subsequently become a triennial event. Koerner also helped to establish a number of national societies in the field, including the Henry Sweet Society for the History of Linguistic Ideas in the UK and the North American Association for the History of the Language Sciences.

Koerner's other editorial enterprises would have wide-ranging positive repercussions for the development of linguistics in the last decades of the 20th century. His founding in 1984 of the journal *Diachronica* would help revitalize historical linguistics at a time when it appeared to be in a terminal decline, and the series *Current issues in linguistic theory*, which he started with Benjamins (whose move from antiquarian book dealer to publisher occurred at Koerner's prompting), offered a stable, high-quality publication outlet for nonmainstream approaches to the study of language at a point when English language academic presses had begun focusing narrowly on Chomskyan generativism. Linguistics would be much less diverse a field today had it not been for these undertakings.

See also: Historical and Comparative Linguistics in the 19th Century; Historiography of Linguistics; Neogrammarians; Saussure, Ferdinand (-Mongin) de (1857–1913).

Bibliography

Cowan W & Foster M K (eds.) (1989). *E. F. Konrad Koerner bibliography.* Bloomington, IN: Eurolingua.

Embleton S, Joseph J E & Niederehe H-J (eds.) (1999). *The emergence of the modern language sciences: studies on the transition from historical-comparative to structural linguistics, in honour of E. F. K. Koerner* (2 vols). Amsterdam and Philadelphia: John Benjamins.

Kirk N & Sidwell P (eds.) (1998). *Practicing Koernerian linguistics: A selection of papers and reviews presented in honour of Professor E. F. K. Koerner.* Melbourne: Association for the History of Language.

Swiggers P (1999). *E. F. K. Koerner: a bio-bibliographical notice.* Leuven: Peeters.

Kono, Rokuro (1912–1998)

T Maruyama, Nanzan University, Nagoya, Japan

Rokuro Kono, a member of the Japan Academy, passed away in 1998 at the age of 85, leaving a legacy of various linguistic achievements. Born in Kobe in 1912, he graduated from the department of linguistics of the University of Tokyo in 1937 and started his academic career in Seoul. After returning to Japan, he spent most of his time at Tokyo University of Education, where he raised several promising academic successors. He assumed various important posts at the Toyo Bunko and Oriental Library and Research Institute for Languages and Cultures of Asia and Africa. He was also a director at the Center for East Asian Cultural Studies for Unesco from 1979 to 1981. He was commended officially as a person of cultural merits in 1993.

His academic achievements are mainly in the following three areas: the study of the Korean language, the study of Chinese historical phonology, and grammatology.

Kono studied the development of grammar and the accentual system from 15th- and 16th-century Korean to Modern Korean. He also made remarkable achievements in the comparative study of Egyptian hieroglyphs, cuneiform characters, and Chinese characters. His interpretation concerning *zhuanzhu* (*chuan-chu*), one type of the six different patterns of development of Chinese characters, is rated highly in the academic world. Although he was a renowned expert in Oriental languages, he also amassed considerable knowledge of classical Greek, Latin, Sanskrit, and Hebrew, not to mention modern languages such as English, German (German, Standard), French, Russian, and, of course contemporary and classical Chinese.

Kono also had a unique and significant approach to the typology of language. Although his ideas concerning typology of language have not been presented in a unified body of writings, they still have had a strong influence on later generations of linguists in Japan. The typological distinctions he raised between double-branch languages, those always requiring a grammatical subject in their sentence structure, like English, and single-branch languages, those not requiring a grammatical subject, like Japanese, are very thought provoking and may contribute much to the future typology of language. His theory was probably the first unique approach to typology born from Japanese cultural and linguistic soil.

Kono was one of the chief editors of *The Sanseido encyclopaedia of linguistics* (6 vols), published from 1988 to 1996 in Japan. It contains descriptions of more than 3500 different languages of the world and about 1500 important terminologies of linguistics.

See also: Chinese; Early Historical and Comparative Studies; Japan: History of Linguistics; Japanese; Korean.

Bibliography

Kono R (1969). 'The Chinese writing and its influences on the scripts of the neighboring peoples: with special reference to Korea and Japan.' *Memoirs of the Research Department of the Toyo Bunko 27,* 83–140.

Kono R (1950). 'On the intensive stem of Middle Korean.' *Gengo Kenkyu 16,* 116–125.

Kono R (1957). 'On the 'tense' system of Middle Korean verbs.' *Memoirs of the Research Department of the Toyo Bunko 16,* 161–182.

Kono R (1977). 'On Chuan-chu.' *Memoirs of the Research Department of the Toyo Bunko 35,* 183–203.

Kono R (1987). 'The Bilingualism of the Paekche Language.' *Memoirs of the Research Department of the Toyo Bunko 45,* 75–86.

Kono R (1979–1980). *Kono Rokuro Chosakushu (selected writings by Rokuro Kono)* (3 vols). Tokyo: Heibonsha.

Kordofanian Languages

T C Schadeberg

Kordofanian is the name of an African language family. It derives its name from that of a former Islamic state with El Obeid as its center.

All 20-odd Kordofanian languages are spoken in the Nuba Mountains in the Republic of the Sudan. The total number of speakers is estimated at around 200 000. Available information about Kordofanian languages is sketchy; no Kordofanian language has been well documented and analyzed. The unity of the Kordofanian language family was first postulated by J. H. Greenberg in 1950; later (in 1963), he classified

Kono, Rokuro (1912–1998)

T Maruyama, Nanzan University, Nagoya, Japan

Rokuro Kono, a member of the Japan Academy, passed away in 1998 at the age of 85, leaving a legacy of various linguistic achievements. Born in Kobe in 1912, he graduated from the department of linguistics of the University of Tokyo in 1937 and started his academic career in Seoul. After returning to Japan, he spent most of his time at Tokyo University of Education, where he raised several promising academic successors. He assumed various important posts at the Toyo Bunko and Oriental Library and Research Institute for Languages and Cultures of Asia and Africa. He was also a director at the Center for East Asian Cultural Studies for Unesco from 1979 to 1981. He was commended officially as a person of cultural merits in 1993.

His academic achievements are mainly in the following three areas: the study of the Korean language, the study of Chinese historical phonology, and grammatology.

Kono studied the development of grammar and the accentual system from 15th- and 16th-century Korean to Modern Korean. He also made remarkable achievements in the comparative study of Egyptian hieroglyphs, cuneiform characters, and Chinese characters. His interpretation concerning *zhuanzhu* (*chuan-chu*), one type of the six different patterns of development of Chinese characters, is rated highly in the academic world. Although he was a renowned expert in Oriental languages, he also amassed considerable knowledge of classical Greek, Latin, Sanskrit, and Hebrew, not to mention modern languages such as English, German (German, Standard), French, Russian, and, of course contemporary and classical Chinese.

Kono also had a unique and significant approach to the typology of language. Although his ideas concerning typology of language have not been presented in a unified body of writings, they still have had a strong influence on later generations of linguists in Japan. The typological distinctions he raised between double-branch languages, those always requiring a grammatical subject in their sentence structure, like English, and single-branch languages, those not requiring a grammatical subject, like Japanese, are very thought provoking and may contribute much to the future typology of language. His theory was probably the first unique approach to typology born from Japanese cultural and linguistic soil.

Kono was one of the chief editors of *The Sanseido encyclopaedia of linguistics* (6 vols), published from 1988 to 1996 in Japan. It contains descriptions of more than 3500 different languages of the world and about 1500 important terminologies of linguistics.

See also: Chinese; Early Historical and Comparative Studies; Japan: History of Linguistics; Japanese; Korean.

Bibliography

Kono R (1969). 'The Chinese writing and its influences on the scripts of the neighboring peoples: with special reference to Korea and Japan.' *Memoirs of the Research Department of the Toyo Bunko 27,* 83–140.

Kono R (1950). 'On the intensive stem of Middle Korean.' *Gengo Kenkyu 16,* 116–125.

Kono R (1957). 'On the 'tense' system of Middle Korean verbs.' *Memoirs of the Research Department of the Toyo Bunko 16,* 161–182.

Kono R (1977). 'On Chuan-chu.' *Memoirs of the Research Department of the Toyo Bunko 35,* 183–203.

Kono R (1987). 'The Bilingualism of the Paekche Language.' *Memoirs of the Research Department of the Toyo Bunko 45,* 75–86.

Kono R (1979–1980). *Kono Rokuro Chosakushu (selected writings by Rokuro Kono)* (3 vols). Tokyo: Heibonsha.

Kordofanian Languages

T C Schadeberg

Kordofanian is the name of an African language family. It derives its name from that of a former Islamic state with El Obeid as its center.

All 20-odd Kordofanian languages are spoken in the Nuba Mountains in the Republic of the Sudan. The total number of speakers is estimated at around 200 000. Available information about Kordofanian languages is sketchy; no Kordofanian language has been well documented and analyzed. The unity of the Kordofanian language family was first postulated by J. H. Greenberg in 1950; later (in 1963), he classified

Kordofanian as a primary branch of his Niger–Congo family.

There are four branches of Kordofanian, named after centrally located towns (names of individual languages are given in parentheses): Heiban (Moro, Tiro, Shirumba, Utoro, Ebang, Laru, Logol, Rere, Warnang, Ko), Talodi (Ngile, Dengebu, Tocho, Jomang, Nding, Tegem), Rashad (Tagoy and Tegali dialect clusters), and Katla (Kalak, Lomorik). Data from wordlists and short grammatical descriptions make it clear that at least the first three branches of Kordofanian have a noun class system (marked by prefixes), which may be taken as evidence for genetic relationship (i.e., common origin) with the large Niger–Congo language family. Lexical evidence for this relationship remained very limited in the early 1990s.

The nine languages of the Kadugli group (Yega, Mudo, Talla, Miri, Tolubi, Kufo, Sangali, Krongo, Talassa) are spoken by about 100 000 people living on the hills lining the southern edge of the Nuba Mountains. One of these languages has been described in a monograph (Reh, 1985). Greenberg originally classified the Kadugli languages as being part of Kordofanian, but this view has since been challenged. The Kadugli languages have systems of nominal classification that distinguish three or four genders, but any existing typological or substantial similarities with Niger–Congo are not sufficient to claim genetic relationship. It seems most likely that Kadugli belongs to the Nilo–Saharan language family – just as do all the other languages that surround Kordofanian.

See also: Niger-Congo Languages; Sudan: Language Situation.

Bibliography

Greenberg J H (1963). *The languages of Africa*. The Hague: Mouton.

Reh M (1985). *Die Krongo-Sprache (niino mo-di): Beschreibung, Texte, Wörterverzeichnis*. Berlin: Dietrich Reime.

Schadeberg T C (1989). 'Kordofanian.' In Bendor-Samuel J (ed.) *The Niger–Congo languages*. Lanham, MD: University Press of America.

Korea, Democratic People's Republic of: Language Situation

R King, University of British Columbia, Vancouver, BC, Canada

The national language of the Democratic People's Republic of Korea is Korean, spoken by approximately 23 million people. North Korea is highly homogeneous: it has no significant ethnic or linguistic minorities, and no official language besides Korean.

The standard language of North Korea, called Munhwaŏ, or 'the Cultured Language,' in order to distinguish it from the standard in South Korea. It claims to be based on the speech of the capital, P'yŏngyang, and to be "a composite of superior linguistic elements of our national language as created and cultivated by the joint efforts of all the people in the northern half of the republic, and has also developed by absorbing those fine elements of the national language as used traditionally everywhere in the south, including Seoul." The six major regional dialects of Korean are all mutually intelligible, with perhaps the exception of the Cheju Island dialect, which is now moribund, but which survives in diasporic enclaves in Japan. The most important foreign language in North Korean education appears to be English, followed by Chinese.

The literacy rate in North Korea is high, usually estimated at 93% or more. Unlike South Korea, in 1949 North Korea abolished the use of Chinese characters in the public media and, also unlike its neighbor, North Korea appears to have taught Chinese characters consistently (and in greater numbers) in the schools since at least the 1950s. Another contrast with South Korea is North Korea's ambitious, elaborate, consistent, and longstanding official attention to issues of language, such as questions of language policy, language planning, and neologisms.

There is growing concern in both North and South Korea about linguistic divergence between the official languages of the north and the south, but in official North Korean publications this is expressed through ridicule and disdain for the pollution of Korean in the south by English (and, to a lesser extent, Japanese).

See also: Korean; Korea, Republic of (South): Language Situation.

Kordofanian as a primary branch of his Niger–Congo family.

There are four branches of Kordofanian, named after centrally located towns (names of individual languages are given in parentheses): Heiban (Moro, Tiro, Shirumba, Utoro, Ebang, Laru, Logol, Rere, Warnang, Ko), Talodi (Ngile, Dengebu, Tocho, Jomang, Nding, Tegem), Rashad (Tagoy and Tegali dialect clusters), and Katla (Kalak, Lomorik). Data from wordlists and short grammatical descriptions make it clear that at least the first three branches of Kordofanian have a noun class system (marked by prefixes), which may be taken as evidence for genetic relationship (i.e., common origin) with the large Niger–Congo language family. Lexical evidence for this relationship remained very limited in the early 1990s.

The nine languages of the Kadugli group (Yega, Mudo, Talla, Miri, Tolubi, Kufo, Sangali, Krongo, Talassa) are spoken by about 100 000 people living on the hills lining the southern edge of the Nuba Mountains. One of these languages has been described in a monograph (Reh, 1985). Greenberg originally classified the Kadugli languages as being part of Kordofanian, but this view has since been challenged. The Kadugli languages have systems of nominal classification that distinguish three or four genders, but any existing typological or substantial similarities with Niger–Congo are not sufficient to claim genetic relationship. It seems most likely that Kadugli belongs to the Nilo–Saharan language family – just as do all the other languages that surround Kordofanian.

See also: Niger-Congo Languages; Sudan: Language Situation.

Bibliography

Greenberg J H (1963). *The languages of Africa.* The Hague: Mouton.

Reh M (1985). *Die Krongo-Sprache (niino mo-di): Beschreibung, Texte, Wörterverzeichnis.* Berlin: Dietrich Reime.

Schadeberg T C (1989). 'Kordofanian.' In Bendor-Samuel J (ed.) *The Niger–Congo languages.* Lanham, MD: University Press of America.

Korea, Democratic People's Republic of: Language Situation

R King, University of British Columbia, Vancouver, BC, Canada

The national language of the Democratic People's Republic of Korea is Korean, spoken by approximately 23 million people. North Korea is highly homogeneous: it has no significant ethnic or linguistic minorities, and no official language besides Korean.

The standard language of North Korea, called Munhwaŏ, or 'the Cultured Language,' in order to distinguish it from the standard in South Korea. It claims to be based on the speech of the capital, P'yŏngyang, and to be "a composite of superior linguistic elements of our national language as created and cultivated by the joint efforts of all the people in the northern half of the republic, and has also developed by absorbing those fine elements of the national language as used traditionally everywhere in the south, including Seoul." The six major regional dialects of Korean are all mutually intelligible, with perhaps the exception of the Cheju Island dialect, which is now moribund, but which survives in diasporic enclaves in Japan. The most important foreign language in North Korean education appears to be English, followed by Chinese.

The literacy rate in North Korea is high, usually estimated at 93% or more. Unlike South Korea, in 1949 North Korea abolished the use of Chinese characters in the public media and, also unlike its neighbor, North Korea appears to have taught Chinese characters consistently (and in greater numbers) in the schools since at least the 1950s. Another contrast with South Korea is North Korea's ambitious, elaborate, consistent, and longstanding official attention to issues of language, such as questions of language policy, language planning, and neologisms.

There is growing concern in both North and South Korea about linguistic divergence between the official languages of the north and the south, but in official North Korean publications this is expressed through ridicule and disdain for the pollution of Korean in the south by English (and, to a lesser extent, Japanese).

See also: Korean; Korea, Republic of (South): Language Situation.

Bibliography

King R (1996). 'Language, politics and ideology in the post-war Koreas.' In McCann D R (ed.) *Korea Briefing.* Asia Society. 109–144.

King R (2005). 'Language and national identity in the Koreas.' In Simpson A (ed.) *Language and national identity in Asia.* Oxford: Oxford University Press.

Kumatani A (1990). Language policies in North Korea.' *International Journal of the Sociology of Language 82,* 88–107.

Korea, Republic of (South): Language Situation

R King, University of British Columbia, Vancouver, BC, Canada

The national language of the Republic of Korea is Korean, spoken by approximately 48 million people. South Korea, like North Korea, is highly homogeneous, but in the last decade some heterogeneity has been introduced in the form of foreign guest workers, mostly from South- and Southeast Asia. There is also a growing number of foreign English language instructors, primarily from Canada, the United States, and Australia. Nonetheless, South Korea has no significant ethnic or linguistic minorities, nor any official language besides Korean. However, there have been controversial calls in recent years to make English a second official language.

The standard language of South Korea claims to be based on the speech of educated persons from the capital, Seoul. The six major regional dialects of Korean are all mutually intelligible, with perhaps the exception of the Cheju Island dialect, which is now moribund, but which also survives in diasporic enclaves in Japan. Unequivocally, the most important foreign language in South Korean education is English. However, the number of South Koreans studying Chinese and Japanese has soared in the past decade, while French and German have declined significantly in popularity. The literacy rate in South Korea is high, usually estimated at 93 percent or more.

Diglossia is not a feature of the South Korean language situation, but most South Koreans still learn approximately 1800 Chinese characters by the time they leave high school. These characters are often mixed in writing with the indigenous Korean script, *Han'gŭl*, in newspapers, journals, and academic writing. The question of whether to abolish the teaching and writing of Chinese characters has been an issue of debate in the South Korean educational and academic establishment for over fifty years, and it continues to be controversial.

Debates on language standardization and policy in South Korea have tended to be the business of rival academic cliques. Government policy (on, for example, orthography and romanization) as implemented by the ministries of education and culture has been inconsistent and, at times, controversial. The founding in 1991 of the National Academy of the Korean Language under the auspices of the Ministry of Culture and Sports, has already done much to rationalize South Korean language policy. At the time of writing, the National Academy of Korean Language is moving cautiously toward joint research projects with its counterpart in North Korea.

There is growing concern in both North and South Korea about linguistic divergence between the official languages of the north and the south. In South Korea this concern is accompanied by acute anxiety over the future viability of Korean under conditions of globalization that favor massive importation of English loanwords and grammatical patterns. Also noteworthy in the sociolinguistic situation in South Korea is internet language: that nation has one of the highest internet connectivity rates in the world, and computer-mediated communication is already having a visible effect on South Korean sociolinguistic life.

See also: Korean; Korea, Democratic People's Republic of: Language Situation.

Bibliography

King R (1996). 'Language politics and ideology in the post-war Koreas.' In McCann D R (ed.) *Korea Briefing.* Asia Society. 109–144.

King R (2005). 'Language and national identity in the Koreas.' In Simpson A (ed.) *Language and national identity in Asia.* Oxford: Oxford University Press.

Kumatani A (1990). 'Language policies in North Korea.' *International Journal of the Sociology of Language 82,* 88–107.

Bibliography

King R (1996). 'Language, politics and ideology in the post-war Koreas.' In McCann D R (ed.) *Korea Briefing.* Asia Society. 109–144.

King R (2005). 'Language and national identity in the Koreas.' In Simpson A (ed.) *Language and national identity in Asia.* Oxford: Oxford University Press.

Kumatani A (1990). Language policies in North Korea.' *International Journal of the Sociology of Language 82,* 88–107.

Korea, Republic of (South): Language Situation

R King, University of British Columbia, Vancouver, BC, Canada

The national language of the Republic of Korea is Korean, spoken by approximately 48 million people. South Korea, like North Korea, is highly homogeneous, but in the last decade some heterogeneity has been introduced in the form of foreign guest workers, mostly from South- and Southeast Asia. There is also a growing number of foreign English language instructors, primarily from Canada, the United States, and Australia. Nonetheless, South Korea has no significant ethnic or linguistic minorities, nor any official language besides Korean. However, there have been controversial calls in recent years to make English a second official language.

The standard language of South Korea claims to be based on the speech of educated persons from the capital, Seoul. The six major regional dialects of Korean are all mutually intelligible, with perhaps the exception of the Cheju Island dialect, which is now moribund, but which also survives in diasporic enclaves in Japan. Unequivocally, the most important foreign language in South Korean education is English. However, the number of South Koreans studying Chinese and Japanese has soared in the past decade, while French and German have declined significantly in popularity. The literacy rate in South Korea is high, usually estimated at 93 percent or more.

Diglossia is not a feature of the South Korean language situation, but most South Koreans still learn approximately 1800 Chinese characters by the time they leave high school. These characters are often mixed in writing with the indigenous Korean script, *Han'gŭl*, in newspapers, journals, and academic writing. The question of whether to abolish the teaching and writing of Chinese characters has been an issue of debate in the South Korean educational and academic establishment for over fifty years, and it continues to be controversial.

Debates on language standardization and policy in South Korea have tended to be the business of rival academic cliques. Government policy (on, for example, orthography and romanization) as implemented by the ministries of education and culture has been inconsistent and, at times, controversial. The founding in 1991 of the National Academy of the Korean Language under the auspices of the Ministry of Culture and Sports, has already done much to rationalize South Korean language policy. At the time of writing, the National Academy of Korean Language is moving cautiously toward joint research projects with its counterpart in North Korea.

There is growing concern in both North and South Korea about linguistic divergence between the official languages of the north and the south. In South Korea this concern is accompanied by acute anxiety over the future viability of Korean under conditions of globalization that favor massive importation of English loanwords and grammatical patterns. Also noteworthy in the sociolinguistic situation in South Korea is internet language: that nation has one of the highest internet connectivity rates in the world, and computer-mediated communication is already having a visible effect on South Korean sociolinguistic life.

See also: Korean; Korea, Democratic People's Republic of: Language Situation.

Bibliography

King R (1996). 'Language politics and ideology in the post-war Koreas.' In McCann D R (ed.) *Korea Briefing.* Asia Society. 109–144.

King R (2005). 'Language and national identity in the Koreas.' In Simpson A (ed.) *Language and national identity in Asia.* Oxford: Oxford University Press.

Kumatani A (1990). 'Language policies in North Korea.' *International Journal of the Sociology of Language 82,* 88–107.

Korean

J H Grayson, University of Sheffield, Sheffield, UK

Phonology

Basic Structures

The Korean language is a nontonal, polysyllabic, agglutinative language belonging to the Altaic family and probably closely related to the Manchu and Tungus members of that language family. The only major modern language to which Korean would appear to be related is Japanese, but the two languages, although similar in most respects grammatically, are significantly different phonologically. Korean and Japanese are therefore linguistic isolates due to the lack of sources to demonstrate the precise linguistic connections between them and with other members of the Altaic family.

Unlike Chinese, Korean lacks true tonal sounds, although it does have vowel stress. The morphological structures of Korean are extremely complex. Korean vocabulary items are built up of multiple morphemes into a highly polysyllabic composition. Like all members of the Altaic family of languages, Korean uses certain morphemes as functional markers to indicate the role of a word within the sentence, as well as mood, tense, location, and the social relationship between the speaker, listener, and the person spoken about.

Triple Consonantal Structure

The consonants of the Korean language are unusual for the triple distinction that is made between soft consonants (lenis consonants), hard, unaspirated consonants, and hard, aspirated consonants. The consonants of the lenis series are *k, n, t, l, m, p, s,* and *ch.* The hard, unaspirated consonants are *kk, tt, pp, ss,* and *tch.* The hard, aspirated consonants are *k', t', p', ch',* and *h*. (These transcriptions follow the orthographic conventions of the McCune-Reischauer System of Romanization, the standard system of scholarly transcription.) The sound of *l* becomes a strongly flapped *r* when placed in an intervowel context. Usually consonants of any of the three series of consonants are pronounced as voiceless, with the exception that the soft consonants *k, t, p,* and *ch* are pronounced *g, d, b,* and *j* when they occur between voiced sounds.

Consonantal Position

A principal phonological feature of Korean is the extreme restriction of consonant position within a given morpheme. Certain sound sequences within a morpheme are not permitted, such as *s* combined with *k*, although the reverse may occur when *k* is final in the preceding morpheme and *s* begins the succeeding morpheme. Certain consonants when in the final position in the morpheme become a strongly dentalized sound. Thus, *t, tt* fortis, unaspirated *t*, *t'* (fortis, aspired *t), s, ch,* and *ch,* and *ch'* (fortis aspirated *ch*) all are pronounced as if they were *t* when they occur in the final position.

Intermorphemic Sound Change

Sound change between syllables is an important feature of the pronunciation of Korean morphemes. This feature, also true of Japanese, is made more difficult for the reader of Korean because it is an orthographic convention that the shape of the individual syllable (morpheme) should be preserved. Although the Korean alphabet itself is highly phonetic, the orthographic convention to preserve the written appearance of the syllable means that the reader must learn a large number of standardized sound changes that occur in the intersyllable position.

Intermorphemic Sound Movement

Sound movement between syllables also occurs. When a syllable that ends in a consonant is followed by a syllable beginning with a vowel, the final consonantal sound passes over to the next syllable. This passage of sound is not represented orthographically.

Nonclustering of Initial Consonants

Clusters of consonants at the beginning of a syllable are not characteristic, there being no equivalents of English *sk, st, str, sh,* and so on.

Triple Vowel System

The vowel system of Korean is as complex as the system of consonants. There are three ranges of vowels: standard vowels (monophthongs), vowel sounds beginning with *y* (rising diphthongs), and a wide range of full diphthongs and diphthongs beginning with the sound *w.* The basic vowels of the monophthong series are pronounced similarly to the vowels of the Romance languages. The monophthongs are *a*, *ŏ, o, u*, and *ŭ*. The *y* series of rising diphthongs are *ya*, *yŏ*, *yo, yu*. The principal diphthongs number 11, although other combinations are possible. Vowels are characterized by phonemic length, which refers to an alteration in tonal height. There is evidence of an earlier stage of vowel harmony that exists as a residual characteristic in certain linguistic contexts.

Grammar and Syntax

General Features

Korean is an agglutinative language with strong elements of fusion and analytical development. The morphological development of word derivation is a well-developed feature of the grammar of the language. Nouns possess a wealth of case forms, possess the grammatical category of specification, and do not possess grammatical gender. There are forms of demonstrative pronouns that indicate varying degrees of spatial relationship.

Number

There are two types of numeral systems: the indigenous Korean system, and the Sino-Korean system that was borrowed as an entire loanword system. Along with the numeral system, there is a system of classifiers that are bound morphemes used as counting words to refer to objects, animals, or people.

Syntactical Markers

The predicatives, verbs, and adjectives of Korean do not have person, number, or gender. They do possess markers indicating social status, tense, and a sentence conclusion. There are three major bands of social status or reference that can be indicated with the special markers, each band containing within it possibilities for further refinements to indicate the precise degree of social relationship existing between the speaker, the listener, or the person spoken about. Tense markers indicate three broad classes of time: the present, the past, and the future (more properly, supposition about the occurrence of an event). Sentence conclusion markers indicate a wide range of moods and meaning, including simple declaration, interrogation, request, demand, suggestion, and reflection. In addition, there are quotative constructions that may be added to the verb to indicate the quotation of a declaration, interrogation, demand, or request. The structure of the verb is verb stem + honorific infix + tense infix + sentence conclusion marker (vs + hi + ti + scm). There is a separate lexical form of the verb that is used to place the verb in alphabetical order in dictionaries, lexicons, and word lists.

Word Order and Word Relation

Word order in sentences for both independent and dependent clauses is always in the sequence of subject, object, predicate. Modifiers, whether of the adjectival or adverbial type, are always in the preposition modifying the word to which they refer. Syntactic relations between words may be expressed by postpositional markers, particles, syntactic nouns, adverbial particles, participles, and the infinitive form of predicatives. Thus, a sentence may consist of a series of clauses, such as an extended adjectival clause modifying a noun, which contains its own subject, object, and predicative with tense and honorific markers attached.

Speech Levels and Honorifics

As a key characteristic of the use of the Korean language is an appropriate use of the system of honorifics that show deference, a sentence must take into consideration three dimensions of speech relationship: (a) the nature of the relationship between the speaker and the listener, (b) the nature of the relationship between the speaker and the person spoken about, and (c) the appropriate way to speak of or about oneself. Any complete sentence will take into account at least one of these dimensions in addition to considerations of tense and mood.

Pronouns, especially for the second person, are very seldom used, the subject of the sentence being understood from the linguistic context.

Sentence Linkage

In sentences containing two independent clauses, the two clauses are linked together through a connection marker attached to the predicative of the first clause. The predicate of the final independent clause will contain markers for honorifics, tense, mood, and sentence conclusion. Where the initial clause is dependent, the connection marker attached to the predicative will indicate the precise relationship of the dependent clause to the independent and principal clause.

Vocabulary

Lexicon

Korean vocabulary is of three types: indigenous Korean vocabulary, Sinitic vocabulary, and loanwords from European languages. Indigenous Korean vocabulary is highly polysyllabic in structure, a feature that is put to good use in sound imitation. Of the world's languages, Korean is one of the most highly onomatopoetic languages. Sinitic vocabulary consists of three subtypes: (a) uniquely Korean terms created by using Chinese characters, (b) direct loanwords from Chinese, and (c) Sino-Japanese loanwords. Sinitic vocabulary consists of terms that are in both ordinary and learned usage, and constitutes more than half of the entire Korean lexicon. European

languages, particularly English, have contributed a number of words, both to the speech of the ordinary person and to the technical speech of professional persons. French, German, Spanish, and Portuguese have also made small contributions to the vocabulary of Korean. There are, or were, a small number of pure Japanese loanwords, but most of these have fallen into disuse through a movement for the purification of the language.

Word Creation

It is still common to use Chinese characters to create new items of technical vocabulary, for example, *nokhwa-gi,* for 'videotape recorder.' Most Sinitic items of vocabulary enter the language as nouns. By attaching the verb *hada* 'to do' in the postposition of the noun, loanwords of this type may be transformed into verbs. By using one of several constructions, such verbs may then be made into adjective or adverbial constructions. By adding *ki* (carrying a sense of continuous action) or *um/m* (carrying an abstract sense) to a verb stem, a verb may be nominalized. Dictionary definitions are often given using *um/m* attached to the final verb.

Parallel Vocabulary Sets

Throughout the vocabulary of Korean, there exists a parallel set of Korean and Sino-Korean vocabulary. Mention was made earlier of the existence of two systems of counting. This feature carries throughout the entire Korean lexicon. Often, but not exclusively, Sino-Korean words are used to name objects or subjects of discourse, while Korean words have a descriptive function. On some occasions, there is no preference in the use of one or the other type of vocabulary; in other instances, it is a matter of honorific or nonhonorific usage. With regard to time, hours are given in Korean numbers, while minutes are given in Sino-Korean numbers. Again, duration of time (i.e., 'it took 1 hour to go home') is given using Korean numerals.

Notwithstanding the enormous impact that Sinitic vocabulary has had on enriching the vocabulary of the Korean language, there has been virtually no influence on the grammar of Korean, possibly because Chinese and Korean derive from two radically different language families.

The Written Language

Modern Korean may be written in one of two forms, either by using only the Korean alphabet (known as *Han'gul* in the Republic of Korea) or a mixed script of *Han'gul* and Chinese characters. In North Korea, the Korean alphabet is used exclusively. In South Korea, usage varies from context to context. A personal letter may use only the indigenous alphabet, while newspapers, textbooks, and better-quality popular books will use a large number of Chinese characters in the text. The more sophisticated and formal a piece of writing is, the more Chinese characters will be used. The Ministry of Education requires a high school graduate to have mastered between 1700 and 1800 characters, the number of characters expected to be encountered in a daily newspaper.

See also: Altaic Languages; Japanese; Korea, Democratic People's Republic of: Language Situation; Korean Lexicography; Korea, Republic of (South): Language Situation; Korean Script: History and Description.

Bibliography

Baek E-J (1984). *Modern Korean syntax*. Seoul: Jung Min.

Dupont R & Millot J (1984). *Grammaire corénne.* Paris: Léopard d'or.

Lee H H B (1989). *Korean grammar.* Oxford: Oxford University Press.

Li J-M (1985). *Grammaire du coréen: vol.1.* Paris: P. A. F.

Li J-M (1991). *Grammaire du coréen: vol. 2.* Paris: P. A. F.

Martin S E (1954). *Korean morphophonemics.* Baltimore, MD: Linguistic Society of America.

Martin S E (1992). *A reference grammar of Korean: A complete guide to the grammar and history of the Korean language.* Tokyo: Tuttle.

Ramstedt G J ([1939] 1979). *A Korean grammar.* Atlantic Highlands, NJ: Humanities Press.

Sohn H-M (1994). *Korean.* London: Routledge.

Song S C (1988). *Explorations in Korean syntax and semantics.* Berkeley, CA: University of California at Berkeley, Centre for Korean Studies.

Yeon J (2003). *Korean grammatical constructions: their form and meaning.* London: Saffron Books.

Korean Lexicography

J Eu, Otago University, New Zealand

The Early Period

The unique Korean alphabet was invented in 1443. It is similar to the Roman alphabet in that it is phonetic, combining vowels and consonants. However, a substantial portion of the Korean vocabulary is based on Chinese characters. A Chinese-based Korean word has an ideographic Chinese character for each of its syllables. Thus, the Korean vocabulary is divided into Chinese-based words and native Korean words.

Shortly after the invention of the Korean alphabet, pronunciation guides that indicated the Korean sounds for Chinese characters appeared, in particular 訓蒙字會 (*Hunmongjahoe*) (1527), which was the first guide to provide both the meaning and sound of Chinese characters. The earliest known dictionaries as such, called 'word collections' (*eohuijip*), date from the late 17th century. Some of them were bilingual or multilingual, while others were mixtures of bilingual and monolingual elements. The bilingual dictionaries invariably had Chinese headwords and gave one simple Korean equivalent for each headword. The multilingual dictionaries simply added Japanese, Mongolian (Halh Mongolian), or Manchurian equivalents to those Chinese-Korean sets.

The mixed works had Chinese characters and words as the headwords, with a Korean equivalent for each headword. But in this case, this entry structure merely followed the tradition of Chinese priority, and in actuality the Chinese headwords served more like a meaning indicator for the Korean words offered as their equivalents, which were the real focus of these works. They can be considered bilingual because they use Chinese characters, and yet their monolingual character lies in the fact that a large portion of the Chinese words used are Chinese-based Korean words.

These word collections are organized according to meaning, or sometimes by the number of the brushstroke of the Chinese headwords, rather than in alphabetical order. They are called 'translation equivalents dictionaries' (*daeyeok sajeon*), because they provide a simple equivalent rather than a definition or linguistic information. Such dictionaries reflected the Chinese-centeredness of the Korean linguistic community of the time.

Modern Monolingual Dictionaries

The first published monolingual Korean dictionary in the modern sense is Moon Se Young's *Korean dictionary* (*Chosuneo sajeon*) (1938). This was the first dictionary to arrange Korean headwords alphabetically and to give definitions as well as linguistic information. Following this, several other Korean dictionaries were compiled, with increased size, many theoretical and methodological innovations, and specialized purposes.

The most notable recent developments are the publication of the *Yonsei Korean dictionary (YKD)* (1998) by Yonsei University and the *Standard Korean dictionary (SKD)* (1999) by the National Institute of the Korean Language. The YKD is the first corpus-based dictionary of Korean, with 50 000 headwords selected according to corpus frequency. Its focus is on current usage. In contrast, the SKD is the largest Korean dictionary ever produced, with 500 000 headwords, combining the entries of previous dictionaries and including archaic words. The SKD presents citations from works of literature, news, magazines, etc., and provides etymology. Currently, there are lexicography centers at Yonsei University and Korea University, where active research is underway to promote further advances in Korean lexicography.

Modern Bilingual Dictionaries

Bilingual dictionaries with Korean or non-Chinese headwords, arranged in alphabetic order, began to appear in the late 19th century. The foreign languages included French, Russian, and English. These dictionaries were written mainly by Western missionaries to Korea. In 1920 the Japanese colonial government in Korea published a Korean-Japanese dictionary.

Since the liberation of Korea in 1945, the rising importance of English has made Korean-English and English-Korean dictionaries the most prominent bilingual dictionaries in Korea. The first K-E and E-K dictionaries written for Korean learners of English were the *New life English-Korean dictionary* (*Shinseng younghan sajeon*) (1946) and its companion volume, the *New life Korean-English dictionary* (*Shinseng hanyoung sajeon*) (1947) by Ryu Hyeong Gi. Since then, numerous other authors and publishers have entered the market. Currently there are five major dictionary publishers producing and updating the five largest K-E and E-K dictionaries in Korea.

In contrast to monolingual lexicography, however, bilingual lexicography has not yet received serious theoretical and methodological attention in Korea. Bilingual dictionaries are compiled by commercial publishing companies, and their weaknesses have been widely criticized. Recently, however, there have

been some systematic reflections about seeking a new direction in Korean bilingual lexicography (Cho, 2002; Eu, 2004; Lee, 2005).

Computational Lexicography in Korean

Most of the above-mentioned recent dictionaries are available on the Internet, and some are available on CD-ROM. Going beyond a simple replica of paper dictionaries, the government-led 21st-Century Sejong Project has been building (since 1998) a large Korean lexical database to comprehensively serve the needs of Korean language–related research and developments such as natural language processing, lexicography, linguistics, education, and artificial intelligence, especially emphasizing a closer cooperation between linguists and computer scientists.

Machine translation between Korean and Japanese – languages that are very similar – has reached a fairly successful standard. For example, Internet sites where Koreans and Japanese can chat by using machine translation are growing in popularity. On the other hand, machine translation between Korean and English is still at a rather rudimentary level. There are Internet sites that search Korean phrases and sentences for their English equivalents, but their usefulness and accuracy are still quite limited.

See also: Bilingual Lexicography; Chinese Lexicography; Chinese; Corpus Lexicography; Japanese Lexicography; Korean; Lexicography: Overview.

Bibliography

Cho J S (1999). 'The transition of Korean dictionaries (Kueo sajeon Byeoncheonsa).' In *The history of transition of Korean newspapers, broadcasting, language, and writing* (Hanguk shinmun bangsong malgul Byeoncheonsa). Korea Press Center.

Cho M O (2002). A critical analysis of Korean-English dictionaries for productive use by Korean EFL learners. Ph.D. diss., University of Exeter.

Eu J (2004). Semantic theories and the foundation of bilingual lexicography. M.Phil. thesis, University of Birmingham.

Hong J S (1996). 'The compilation of Korean dictionaries– the results and tasks (Koogeo sajeon pyeonchan–ku seonggoawa koaje).' *Hangook Eohak 3(1).*

Lee B G (2000). *The history and direction of Korean lexicography.* (Hanguk sajeonui Yeoksawa Banghyang). Taehaksa.

Lee J H (2005). *A critique of English-Korean dictionaries.* (Younghan sajeon bipan). Kungri.

Nam K I (2005). 'Thirty years of Korean lexicography (Kugeo sajeon Pyeonchanui samshipnyeon).' In *Proceedings of the 21st National Conference of the Korean Language Research Circle.*

Kotarbiński, Tadeusz (1912–1998)

R A Rothstein, University of Massachusetts, Amherst, MA, USA

Tadeusz Marian Kotarbiński was born on March 31, 1886, in Warsaw and died there on October 3, 1981. He studied technical subjects and architecture in Cracow, Lwów (now Lviv, Ukraine) and Darmstadt before taking up philosophy and the classics at Jan Kazimierz University in Lwów, where he earned his doctorate in 1912. He taught Greek and Latin in a Warsaw high school until 1918, when he began teaching at Warsaw University. During the interwar period he was active as a teacher and scholar, and also took part in the public debates of the time, speaking out and writing, for example, against anti-Semitism and clericalism.

During the Nazi occupation of Poland Kotarbiński took part in clandestine teaching; after the war he helped to organize the university in Łódź. He served as its first rector from 1945 to 1949, simultaneously teaching and chairing the Philosophy Department at Warsaw University. He eventually returned to Warsaw, remaining there until his death.

Kotarbiński's main areas of interest were praxiology and ethics, but he also devoted attention to formal logic, logical semantics, and the history of logic. He was the author as well of three volumes of aphoristic poetry.

Kotarbiński called his philosophical approach 'reism' (or 'concretism') and viewed its principles as "declaring war against the hypostases of linguistic origin," the faulty assumption that, for example, properties like 'roundness' or 'blueness' exist because natural languages have words for them. He drew a distinction between genuine names and apparent names. The former are names of persons or things; they include singular, general, and empty names (e.g., *Kotarbiński, philosopher, unicorn*, respectively). The latter are names of properties, relations, events, or states (*truthfulness, equality, departure, health*, respectively). Singular names, which refer to individuals or things, can serve as grammatical subjects.

Korean Script: History and Description

C W Kim, University of Illinois, Urbana, IL, USA

History

Use of Chinese Characters in the Early Period

By the beginning of the Common Era, there were three nations in the Korean peninsula and southern Manchuria. Effective governance of the state required a writing system if only to transmit royal edicts to local administrators, yet Korea had no script of its own. Since Korea had been under the heavy cultural influence of China, it was natural that Korea borrowed Chinese writing (*hanmun* 漢文) and characters (*hancha* 漢字) and adapted them to write Korean. Three such adaptive systems are briefly described below.

Both *idu* (吏讀 clerk reading) and *hyangch'al* (鄕札), which apparently have been in use since the 6th century, are mixed transcription systems in which content words are written with Chinese characters for their meanings, while function words borrow pronunciation only. The former is called *hun* (訓 semantic) reading, and the latter, *ŭm* (音 phonetic) reading. For example, a Korean phrase *po-go* 'see-and' would be transcribed and read as follows:

Chinese writing:	見	古
Meaning in Chinese:	'see'	'old'
Pronunciation in Chinese:	[kjen]	[go]
Intended reading in Korean:	[po]	[go]
Intended meaning in Korean:	'see'	'and'

The major difference between *idu* and *hyangch'al* is that while the former maintains the Chinese word order, the word order in the latter is altered to conform more with the word order in Korean. This process was facilitated with the help of *kugyŏl* (口訣), signs of grammatical markers written between phrases in the text. These signs were highly simplified Chinese characters in *ŭm*-reading denoting affixes and grammatical markers, helping the reader in comprehending the text. For example, *go* (古), a verbal connective seen above, was simplified to 口, just the bottom part; the character for topic marker *ŭn* (隱) was simplified to 阝, the left radical of the letter; a conditional affix *myŏn* (面) was simplified to 丆, just the top part of the letter, while a copula marker *i* (是) was simplified to the last stroke at the bottom, 乀, etc. These letter-signs were necessarily small and simplified as they had to be inserted into a small space between phrases and lines. They helped the reader construct sentential structures of Korean from Chinese writing.

These conventions of rendering Korean in Chinese were quite ill-fit. A 15th-century scholar expressed the effort as trying to fit a square pole into a round hole. The mismatch would disappear with a stroke of genius.

Invention of a Native Script: *Hangul*

Authorship and Motivation In volume 102 of *Sejong sillok* (世宗實錄 Chronicles of King Sejong), it is matter-of-factly written that "In December of the 25th year [of his reign = 1443], His Majesty personally devised ("親制") twenty-eight letters of *ŏnmun* (諺文 vernacular script)." Thus is announced one of truly great inventions in the history of mankind. Three years later in October 1446, it was officially promulgated as *Hunmin chŏng'ŭm* (訓民正音 (The Correct Sounds for the Instruction of the People). The script is now generally known as *hangul* (*han'gŭl*), meaning 'Korean writing'.

Classical Chinese phonology, which Sejong studied and had his scholars consult, divided the syllable into initials and rhymes. Sejong realized, however, that syllable-initial consonants and syllable-final consonants had equal weight as phonological units and thus could be written with the same letters. This tripartite division of the syllable isolated the syllable-nuclei and required Sejong to devise separate vowel letters. A new alphabet was thus born. The King's 28 letters, of which 17 were consonants and 11 were vowels, are given below. Note that vowel letters include diphthongs beginning with a semi-vowel [j].

Consonant letters	Vowel letters
ㄱ [k, g]	ㅏ [a]
ㅋ [k^h]	ㅑ [ja]
ㄴ [n]	ㅓ [ə] (ŏ)
ㄷ [t, d]	ㅕ [jə] (yŏ)
ㅌ [t^h]	ㅗ [o]
ㄹ [ɾ, l]	ㅛ [jo]
ㅁ [m]	ㅜ [u]
ㅂ [p, b]	ㅠ [ju]
ㅍ [p^h]	ㅡ [ɨ] (ŭ)
ㅅ [s]	ㅣ [i]
ㅈ [tʃ, ʤ]	• [ɔ]
ㅊ [$tʃ^h$]	
ㅿ [z]	
ㅇ (zero-initial or [ɦ])	
ㆆ [ʔ]	
ㅎ [h]	
ㆁ [ŋ]	

Principles of Graphic Structure Aside from the fact that this is a truly phonemic alphabet, what is also remarkable is that Sejong devised letters in such a way

that the graphic shapes showed the internal relations among the letters. To cite a few obvious ones:

- The visual distinction between the consonantal and vocalic letters is apparent. Vowels contain long bars, vertical or horizontal, with short strokes (originally dots) stemming from them, while consonants have more complex geometric figures.
- Vowel shapes reflect vowel harmony, which was active in 15th-century Korean. Letters with side strokes (dots originally, e.g., ·ㅣ → ㅓ) upward or rightward are 'bright' (*yang* 陽) vowels, while those with downward or leftward side strokes are 'dark' (*yin* 陰) vowels.
- The 'stroke-adding' (*kahoek* 加 劃) principle is applied to derive occlusives and aspirates (*ch'oryŏ* 稍厲 a little harsh) from base consonants, as can be seen in the chart below. One can also tell at a glance that a consonant is a sibilant, as only sibilants have slanting strokes.

	Base shape	One stroke	Two strokes
Labials	ㅁ [m]	ㅂ [p]	ㅍ [pʰ]
Alveolars	ㄴ [n]	ㄷ [t]	ㅌ [tʰ]
Velars	ㆁ [ŋ]	ㄱ [k]	ㅋ [kʰ]
Sibilants	ㅅ [s]	ㅈ [tʃ]	ㅊ [tʃʰ]
Glottals	ㅇ [ɦ] or Ø	ㆆ [ʔ]	ㅎ [h]
	Continuants	Occlusives	Aspirates

Origin and Sources There are three theories on the possible origin of the letter shapes.

1. A leading theory is that the forms of five basic letters were modeled after their articulatory configurations. This theory derives from a description in *Hunmin chŏng'ŭm haerye* (訓民正音解例 (Explanations and Examples of the Correct Sounds for the Instruction of the People), published in 1446 but only discovered in 1940, where it is stated that these letters 'depict outlines' (상형 象形) of speech organs. The relation between the letter forms and articulatory shapes is clear from the illustrative diagrams in Figure 1.
2. A gaining theory is that the letter forms imitated the Mongolian 'Phags-pa script. It comes from a new interpretation of a much disputed phrase in *Haerye* saying that the letter forms 'imitated *kojŏn*' (방고전 倣古篆). *Kojŏn* was initially thought to refer to 'old seal' characters, one of the stylized forms of old Chinese writing. But Ledyard (1966, 1997) argued that the character 古 (*ko/go*) referred to 蒙古 (*mongo*) or Mongolia.

   When Mongols founded the Yuan dynasty in the 13th century, Emperor Khubilai Khan felt that the new empire needed a new unified script and had a learned monk named 'Phags-pa invent a new script, which is now known as the 'Phags-pa script. There are indeed some uncanny similarities between certain 'Phags-pa and *hangul* letters. For example, from 'Phags-pa letter ꡂ [g], one can derive *hangul* ㄱ by removing the dangling box, or from 'Phags-pa letter ꡙ [ℓ], one can derive *hangul* ㄹ by rotating the letter 90 degrees clockwise and by straightening the middle line. These and other similarities, while not systematic, are a little too striking to relegate them to a simple accident.
3. Yet another theory has it that the letter forms derived from *kugyŏl* signs. To the extent that *kugyŏl* signs evolved from Chinese characters, one can say that this is a theory of Chinese origin. The chief proponent of this theory is W. Kim (1996). Recall that *kugyŏl* as signs of grammatical particles inserted between phrasal structures in Chinese writing came to be highly simplified and stylized, for example, 又 for 奴 (*ro*, Instrumental), 口 for 古 (*go*, Connective), 丿 for 底 (*dyŏ*, Inchoative), 厶 for 矣 (*ŭi*, Genitive), etc. Kim argues that these simplified *kugyŏl* letters eventually evolved into *hangul* letters, much as Japanese *katakana* letters evolved from simplified Chinese characters, e.g., カ [ka] from 加, タ [ta] from 多, セ [se] from 世, リ [ri] from 利, etc.

There have also been suggestions that the Sanskrit (Devanagari) script or the Uighur script served as a model of *Hunmin chŏng'ŭm* letters. But since these scripts were not phonemic alphabets in the strict sense, as the vowel [a] was assumed for every consonant and other vowels were marked as diacritics, it is unlikely that Sejong received the main inspiration from these alpha-syllabaries.

It is also highly unlikely that any one theory is exclusively right. King Sejong and his court scholars were well versed in phonologies and writing systems of the neighboring languages and undoubtedly consulted them as references. Sejong's innovative genius



Figure 1 Letterforms of base *hangul* letters and their articulatory shapes.

is not that he created something out of nothing, but rather that he created something much better in graphic structures, simplicity, and elegance than any writing system known then or now.

Post-15th Century Changes

Hangul, born in a royal crib, nevertheless took a painful and long time to grow. Not unexpectedly, it met fierce opposition in a Sino-centric world. No lesser figure than a Vice Director of the Royal Academy (Ch'oe Manli) stated in a memorial that "disregarding Chinese in favor of a vernacular script would be identifying ourselves with the barbarians." Chinese writing continued to be used until the end of the 19th century in official papers and among literati. *Hangul* survived mainly in the hands of women and commoners.

Among the original letters, ㅿ [z] and ㆆ [ʔ] became disused and lost, the former because the sound itself was lost and the latter because it was ill-fated from the beginning as a prescriptive letter to record orthodox Chinese pronunciation. ㆁ [ŋ], occurring as syllable-final only, was replaced by ㅇ, a null syllable-initial. Geminate letters, ㄲ, ㄸ, ㅆ, ㅉ, which were initially used to record voiced Chinese sounds, were also disused, but were later revived to record tense consonants, replacing the old ways of writing them, i.e., ㅺ, ㅼ, ㅽ, ㅾ. This situation resulted in much confusion, and became the source of a heated debate among Korean scholars as to whether or not these letters actually represented consonant clusters, [sk, st, sp], etc.

Vowel shift triggered the loss of the vowel ㆍ [ɔ], and the eventual decay of vowel harmony. Former diphthongs came to represent newly emerged front vowels, e.g., ㅐ [aj → æ], ㅔ [əj → e].

What was primarily a phonemic orthography at the time of *hangul*'s invention has now become a morphophonemic orthography, following the proposal for Unified Orthography in 1933 by the Korean Language Society. Thus, 새미기픈므른 /sæ.mi.ki.p^{h}ɨn.mɨ.rɨn/, 'the water of a deep well', is now written 샘이 깊은 물은 /sæm.i kip^h.ɨn mul.ɨn/, preserving the morphemic identity (underlying form) of the lexical items and with word space. Today, *hangul* as the script of Korean is used by 70 million in both Koreas, with nearly 100% literacy thanks to *hangul*'s elegant simplicity.

Description

Inventory

King Sejong (r. 1418–1450) named his script *Hunmin chŏng'ŭm* 'Correct sounds for the instruction of people'. In the following centuries when Chinese still functioned as the official writing, it came to be known as *ŏnmun* 'a vulgar script'. Today it is generally known as *hangul* (*han'gŭl*), meaning 'a great writing' or possibly 'the Han (national) script', a coinage by the linguist Chu Sigyŏng (1886–1914). The script is called *urigŭl* 'our script' in North Korea (People's Democratic Republic of Korea).

Modern *hangul* is an alphabetic script consisting of 24 letters, 14 of which are letters for consonants and 10 are letters for vowels. Vowel letters include [j]-initial diphthongs, as can be seen in the following table. Voiced obstruents are intervocalic allophonic variants. ㄹ is [r] syllable-initially but [l] syllable-finally. ㅇ is silent syllable-initially, but is [ŋ] syllable-finally. Although the two had distinct letters, ㅇ and ㆁ, respectively, at the time the script was created, the two merged later, since [ŋ] never occurs syllable-initially in Korean. The order given below is that of dictionary entries in South Korea:

Consonant letters	Vowel letters
ㄱ [k, g]	ㅏ [a]
ㄴ [n]	ㅑ [ja]
ㄷ [t, d]	ㅓ [ə] (ŏ)
ㄹ [r, l]	ㅕ [jə] (yŏ)
ㅁ [m]	ㅗ [o]
ㅂ [p, b]	ㅛ [jo]
ㅅ [s]	ㅜ [u]
ㅇ (zero-initial or [ŋ])	ㅠ [ju]
ㅈ [tʃ, ʤ]	ㅡ [ɨ] (ŭ)
ㅊ [tʃ^h]	ㅣ [i]
ㅋ [k^h]	
ㅌ [t^h]	
ㅍ [p^h]	
ㅎ [h]	

Letters for obstruent consonants may be doubled to create geminates (or tense consonants), and some vowel letters combine to form new vowels including [w]-initial diphthongs, as shown below:

Geminates	Compound vowels
ㄲ [k']	ㅐ [æ] (from ㅏ + ㅣ)
ㄸ [t']	ㅒ [jæ]
ㅃ [p']	ㅔ [e] (from ㅓ + ㅣ)
ㅆ [s']	ㅖ [je]
ㅉ [tʃ']	ㅘ [wa]
	ㅚ [œ] (from ㅗ + ㅣ)
	ㅝ [wə] (wŏ)
	ㅟ [wi]
	ㅢ [ɨj] (ŭi)
	ㅙ [wæ]
	ㅞ [we]

Thus, the total number of letters including basic and complex letters is 30, 19 of which are consonant

letters and 21 are vowel letters. A large number of vowel letters stem from the fact that glides function as part of nuclei vowels in the syllable structure of Korean (Kim and Kim, 1991). There are also the following eleven syllable-final two-consonant clusters: ㄳ [ks], ㄵ [ntʃ], ㄶ [nh], ㄺ [lk], ㄻ [lm], ㄼ [lp], ㄽ [ls], ㄾ [lt^h], ㄿ [lp^h], ㅀ [lh], and ㅄ [ps]. No other possible combinations occur, e.g., ㄱㄹ, ㄴㄷ, ㄷㄹ, ㄹㄹ, ㅂㅍ, ㄹㄲ, ㅗㅓ, ㅜㅏ, ㅠㅏ, ㅛㅑ, ㅠㅕ, ㅠㅣ, ㅛㅣ, ᆖ, etc., although ㄹㄹ and ᆖ, have been proposed by some linguists as letters denoting lateral ㄹㄹ [ℓ] and ᆖ [jɨ], respectively, the latter glide-vowel being the pronunciation of [jə] in some dialects. No diacritic mark, even an apostrophe, is used.

General Characteristics of Script Shapes

Hangul script is quite unique in that its geometric shapes of basic letters are iconic symbols and that they reveal internal structural relations as described below.

1. Basic consonant letters depict their articulatory configurations (e.g., ㄱ depicts the tongue back raised against the velum, while ㄴ depicts the tongue tip making a contact with alveola (see Figure 1). Three basic vowel letters symbolize three elements, i.e., ㅣ 'a (standing) man', ㅡ 'the (flat) earth', and • (now extinct) 'the (round) heaven'.
2. Consonant and vowel shapes are distinguished from each other. The distinguishing stroke in vowels is a bar, either horizontal or vertical, by itself or with at most two short side strokes (originally dots) to the right or left of a vertical bar, or above or below a horizontal bar. Consonants have more complex geometric shapes with angled strokes.
3. In vowels, two short strokes tell that the letter represents a complex vowel [j] + V where V is the same vowel with only one side stroke, e.g., ㅑ [ja] = [j] + ㅏ [a], ㅠ [ju] = [j] + ㅜ [u], ㅒ [jæ] = [j] + ㅐ [æ], etc. [w]-glide is represented with either ㅗ [o] or ㅜ [u] before another vowel, e.g., ㅘ [wa] ([o] + [a]), ㅟ [wi] ([u] + [i]), etc.
4. In consonants, only sibilants have slanting strokes, i.e., ㅅ [s], ㅈ [tʃ], ㅊ [tʃ^h], while non-sibilants have right-angle strokes or circles, e.g., ㄱ [k], ㄷ [t], ㅂ [p], ㅁ [m], ㄹ [l], ㅇ [ŋ].
5. The so-called add-a-stroke principle yields derived sounds in a systematic way. For example, adding a stroke to a basic obstruent makes it an aspirated sound, e.g.:

   ㄱ [k] + stroke = ㅋ [k^h],
   ㄷ [t] + stroke = ㅌ [t^h],
   ㅈ [tʃ] + stroke = ㅊ [tʃ^h],

   In fact, the [j]-glide vowels ([3] above) may be regarded as having been derived from basic vowels via stroke adding.

These features are indeed quite unique and are probably not found in any other script in the world. In the Roman alphabet, for example, there is no internal relation between /p/ and /b/, /t/ and /d/, /k/ and /g/, etc. The subsegmetal features in the Korean script prompted Sampson (1985) and Kim (1988) to say that *hangul* is the only featural script in the world.

Writing Conventions

Unlike in Chinese writing where lines are vertical from right to left, contemporary Korean writing proceeds from left to right in horizontal lines as in English writing. In individual letter writing, *hangul* letters proceed from left to right and top to bottom. In addition, although *hangul* is a fully alphabetic script, it is written in syllable blocks such that one block consists minimally of an initial (onset) consonant and a nucleus vowel. It is because of this requirement that ㅇ is used as a filler letter when there is no initial consonant in the syllable. When the vowel letter is of a vertical bar, C and V are written side by side in this order, e.g., 가 /ka/ 'go'(Root). But if V consists of a horizontal bar, the V is written **below** the C, e.g., 구 /ku/ 'nine'. When there is a final (coda) consonant, this is written at the bottom of the block, regardless of whether the syllable has a vertical or a horizontal vowel, e.g., 간 /kan/ 'went', 국 /kuk/ 'soup', 한글 /han.kul/ '*hangul*', 닭 /talk/ 'chicken'.

Sizes of individual letters are slightly adjusted to fill the syllable block. Thus, ㄱ is elongated in 가 /ka/ but is flattened in 구 /ku/, and both consonants and vowels become slightly smaller when there is a coda, e.g., in 간 /kan/ or 국 /kuk/. This was a substantial problem in the age of typewriters, and some called for linearization or *karo-s'ŭgi* 'side-by-side writing', e.g., ㅎㅏㄴㄱㅡㄹ for 한글 /hankul/, but the advent of word processors quickly silenced the proponents.

Thus, *hangul* is an alphabetic script with subsegmental features but written in syllable blocks. This led Taylor (1979) to call *hangul* an alphabetic syllabary or an alpha-syllabary. This is another advantage of *hangul*. An alphabetic script makes it easier to learn, and a syllabic writing makes it easier to read. An experiment showed that a class of American college students first learning Korean read one and a half times faster in block writing than students in another class who were initially taught to write and read in linearized writing (Kim and Sohn, 1986).

Differences between North and South Korea

In North Korea, the same script is used, but is called *urigul* (*urigŭl*) 'our script', probably just to be different from its southern neighbor, just as the language is known by two different names, *p'yojunŏ* 'the standard language' in the south, but *munhwaŏ* 'the cultured language' in the north.

Both states have adopted the Unified Orthography proposed by the Korean Language Society in 1933, which is basically morphophonemic. Each state has since made minor modifications with varying details. In general the northern variety is more morphemic than the southern.

The order of the alphabet is slightly different in the two Koreas. First of all, in North Korea, ㅇ is the last consonant letter, not the 8th as in South Korea, probably because as the syllable-initial letter, it has a null phonetic value. In addition, characters consisting of two basic letters follow all single letters so that ㄲ, ㅼ, ㅽ, etc., follow ㅇ. But in South Korea, ㄲ immediately follows ㄱ, and ㅼ follows ㄷ, etc. Consonant clusters, however, are treated differently from geminates so that while the order of ㄱ, ㄲ and ㄳ as syllable-finals is as given in South Korea, it is ㄱ, ㄳ, . . . ㄲ in North Korea. That is, while ㄱ is immediately followed by ㄳ, the rest of the single letters intervene between ㄳ and ㄲ. The same principle applies in the order of vowel letters in the north, that is, double-letter characters such as ㅐ, ㅔ, ㅚ follow all single basic letters, and triple-letter characters ㅙ and ㅞ follow all double-letter vowel characters. In South Korea, however, complex letters immediately follow the basic letter from which they derive, e.g., ㅐ (ㅏ + ㅣ) comes after ㅏ but before ㅑ, and ㅘ, ㅙ and ㅚ immediately follow ㅗ, in this order.

Another difference between the two Koreas is the status of Sino-Korean vocabulary and characters (*hancha*). In the north it was determined early that a socialist state had no place for *hancha* because it is the property of the elite class, and the use of *hancha* was abolished altogether. This policy may have had a political motive of facilitating mass indoctrination. In the south, there has been a gradual decline in the use of *hancha*, even as purists and pragmatists heatedly debated the fate of *hancha*. Today about 1800 characters are taught in South Korea, 900 each in middle and high schools. *Hancha* is in fact now seeing a bit of a revival, due to the emergence of China as a powerhouse and due to a recognition that Chinese characters are becoming a lingua franca in East Asia.

See also: China: Writing System; Korean.

Bibliography

Hope E R (1957). 'Letter shapes in the Korean *ŏnmun* and Mongol hP'agspa alphabets.' *Oriens 10,* 150–159.

Kim C W (1978). 'Linguistics and language policies in North Korea.' *Korean Studies 2,* 159–175.

Kim C W (1988). 'On the origin and structure of the Korean alphabet.' In Kim C W (ed.) *Sojourns in Language II.* Seoul: Tower Press. 721–734.

Kim C W (2000). 'The legacy of King Sejong the Great.' In Kim C W (ed.) *Literacy and writing systems in Asia: Studies in the Linguistic Sciences 30,* 2–12.

Kim C W & Kim H Y (1991). 'The *character* of Korean glides.' *Studies in the Linguistic Sciences 21,* 113–125.

Kim C W & Sohn H (1986). 'A phonetic model for reading: evidence from Korean.' *Studies in the Linguistic Sciences 21,* 113–125.

Kim J-P (1983). 'The letterforms of *Han'gul:* its origin and process of transformation.' In Korean National Commission for UNESCO (ed.) *The Korean language.* Seoul: Sisayong'osa. 80–102.

Kim W (1996). *Ŭmun-gwa muncha.* (Phonological segments and characters). Seoul: Sin'gumunhwa-sa.

Kim Y-K (1959). 'Written language of Korean.' *Koreana Quarterly 1,* 121–129.

Kim-Renaud Y-K (ed.) (1997). *The Korean alphabet: its history and structure.* Honolulu: University of Hawai'i Press.

Kim-Renaud Y-K (1997). 'A brief description of the Korean alphabet.' In Kim-Renaud Y-K (ed.) *The Korean alphabet: its history and structure.* Honolulu: University of Hawai'i Press. 279–287.

King R (1996). 'Korean writing.' In Daniels P T & Bright W (eds.) *The world's writing systems.* Oxford: Oxford University Press. 218–227.

Ledyard G (1966). *The Korean language reform of 1446: the origin, the background, and early history of the Korean alphabet.* Ph.D. diss., University of California at Berkeley.

Ledyard G (1997). 'The international linguistic background of the Correct Sounds for the Instruction of the People.' In Kim-Renaud Y-K (ed.) *The Korean alphabet: its history and structure.* Honolulu: University of Hawai'i Press. 31–87.

Lee K-M (1963). English synopsis in *Kugŏ p'yogipŏb-ŭi yŏksajŏk kochal* (A historical study of the Korean writing system). Seoul: Hanguk Yonguwon.

Ramsey S R (1992). 'The Korean alphabet.' In Kim-Renaud Y-K (ed.) *King Sejong the Great.* Washington, DC: International Circle of Korean Linguistics and George Washington University. 41–50.

Sampson G (1985). *Writing systems.* Stanford, CA: Stanford University Press.

Taylor I (1979). 'The Korean writing system: an alphabet? A syllabary? A logography?' In Kolers P A, Wrolstad M E & Bouma H (eds.) *Processing of visible language.* New York: Plenum. 67–82.

been some systematic reflections about seeking a new direction in Korean bilingual lexicography (Cho, 2002; Eu, 2004; Lee, 2005).

Computational Lexicography in Korean

Most of the above-mentioned recent dictionaries are available on the Internet, and some are available on CD-ROM. Going beyond a simple replica of paper dictionaries, the government-led 21st-Century Sejong Project has been building (since 1998) a large Korean lexical database to comprehensively serve the needs of Korean language–related research and developments such as natural language processing, lexicography, linguistics, education, and artificial intelligence, especially emphasizing a closer cooperation between linguists and computer scientists.

Machine translation between Korean and Japanese – languages that are very similar – has reached a fairly successful standard. For example, Internet sites where Koreans and Japanese can chat by using machine translation are growing in popularity. On the other hand, machine translation between Korean and English is still at a rather rudimentary level. There are Internet sites that search Korean phrases and sentences for their English equivalents, but their usefulness and accuracy are still quite limited.

See also: Bilingual Lexicography; Chinese Lexicography; Chinese; Corpus Lexicography; Japanese Lexicography; Korean; Lexicography: Overview.

Bibliography

Cho J S (1999). 'The transition of Korean dictionaries (Kueo sajeon Byeoncheonsa).' In *The history of transition of Korean newspapers, broadcasting, language, and writing* (Hanguk shinmun bangsong malgul Byeoncheonsa). Korea Press Center.

Cho M O (2002). A critical analysis of Korean-English dictionaries for productive use by Korean EFL learners. Ph.D. diss., University of Exeter.

Eu J (2004). Semantic theories and the foundation of bilingual lexicography. M.Phil. thesis, University of Birmingham.

Hong J S (1996). 'The compilation of Korean dictionaries–the results and tasks (Koogeo sajeon pyeonchan–ku seonggoawa koaje).' *Hangook Eohak 3(1).*

Lee B G (2000). *The history and direction of Korean lexicography.* (Hanguk sajeonui Yeoksawa Banghyang). Taehaksa.

Lee J H (2005). *A critique of English-Korean dictionaries.* (Younghan sajeon bipan). Kungri.

Nam K I (2005). 'Thirty years of Korean lexicography (Kugeo sajeon Pyeonchanui samshipnyeon).' In *Proceedings of the 21st National Conference of the Korean Language Research Circle.*

Kotarbiński, Tadeusz (1912–1998)

R A Rothstein, University of Massachusetts, Amherst, MA, USA

Tadeusz Marian Kotarbiński was born on March 31, 1886, in Warsaw and died there on October 3, 1981. He studied technical subjects and architecture in Cracow, Lwów (now Lviv, Ukraine) and Darmstadt before taking up philosophy and the classics at Jan Kazimierz University in Lwów, where he earned his doctorate in 1912. He taught Greek and Latin in a Warsaw high school until 1918, when he began teaching at Warsaw University. During the interwar period he was active as a teacher and scholar, and also took part in the public debates of the time, speaking out and writing, for example, against anti-Semitism and clericalism.

During the Nazi occupation of Poland Kotarbiński took part in clandestine teaching; after the war he helped to organize the university in Łódź. He served as its first rector from 1945 to 1949, simultaneously teaching and chairing the Philosophy Department at Warsaw University. He eventually returned to Warsaw, remaining there until his death.

Kotarbiński's main areas of interest were praxiology and ethics, but he also devoted attention to formal logic, logical semantics, and the history of logic. He was the author as well of three volumes of aphoristic poetry.

Kotarbiński called his philosophical approach 'reism' (or 'concretism') and viewed its principles as "declaring war against the hypostases of linguistic origin," the faulty assumption that, for example, properties like 'roundness' or 'blueness' exist because natural languages have words for them. He drew a distinction between genuine names and apparent names. The former are names of persons or things; they include singular, general, and empty names (e.g., *Kotarbiński, philosopher, unicorn*, respectively). The latter are names of properties, relations, events, or states (*truthfulness, equality, departure, health*, respectively). Singular names, which refer to individuals or things, can serve as grammatical subjects.

General names (of people or things) can serve as predicates, or as subjects in universal sentences. Empty names have no denotation, but can be replaced by synonymous combinations of singular and general terms. An apparent name, Kotarbiński argued, can only legitimately be used if the expression in which it occurs can be translated into an expression containing only genuine names. Thus we can get away with saying *John's anger was shortlived* because we can translate it into *John was not angry for long.* Saying the former no more commits us to the existence of 'anger' than saying *Mary took umbrage at my words* commits us to the existence of 'umbrage'; indeed, all that exist are objects (including people).

Kotarbiński was one of the first philosophers to consider the possibility of a three-valued logic (before Jan Łukasiewicz or Emil Post in the 1920s). He came to this position in 1912 from consideration of the possibility of creative activity. Assuming that truth is eternal, and that a statement S affirming the existence of an object O can be true if and only if O exists, it should be impossible to create O. For until O is created, S cannot be true, therefore – by one version of the law of the excluded middle – S must be (eternally) false and therefore O's existence cannot be affirmed. Kotarbiński rejected this version of the Aristotelian principle, suggesting instead that before O is created, S is neither true nor false, but indefinite.

Bibliography

Kotarbiński T (1965). *Praxiology: an introduction to the science of efficient action.* Warsaw and Oxford: PWN-Pergamon Press.

Kotarbiński T (1966). *Gnosiology: the scientific approach to the theory of knowledge.* Warsaw and Oxford: PWN-Pergamon Press.

Kotarbiński T (1968). 'The problem of the existence of the future.' *The Polish Review 13,* 7–22.

Woleński J (ed.) (1990). *Kotarbiński: logic, semantics, and ontology.* Dordrecht and Boston: Kluwer Academic Publishers.

Koya:ng *See:* Gondi.

Krapf, Johann Ludwig (1810–1881)

S Pugach, Ohio State University, Lima, Ohio

Johann Krapf was born on February 11, 1810, in Derendingen, Württemburg, and died on November 26, 1881, in Korntal, Württemburg. One of four children, he was the son of Johann Jakob Krapf, a farmer, and Katherina Maria Braun. He attended the Osterbergschule in Tübingen, where he discovered a talent for languages, studying Greek, Latin, French, Italian, and Hebrew (Pirouet, 1999). Krapf's next step was training to become a missionary at the Basel Missionary Institute, a training he interrupted to study theology at the University of Tübingen. In 1834 Krapf returned to the mission after a five-year hiatus, in which time he had been ordained and become a parish minister.

At this point the German Krapf joined the British Church Missionary Society, and in 1837 was sent to help found a mission station in Abyssinia. After unsuccessfully trying to establish the station and being driven from the region, Krapf moved to Shoa, in southern Ethiopia, before he was ultimately forced to go even further south, to Mombasa (Miehe, 1983). During his 19 years in Africa, Krapf achieved fame in two arenas: as a writer of historical travelogues, and as a linguist who compiled grammars and biblical translations for several languages. In this second capacity, Krapf belonged to a group of German missionaries, among them S. W. Koelle and J. F. Schön, who introduced the discipline of African linguistics to Europe (*see* **Koelle, Sigismund Wilhelm (1823–1902)**).

Once in Ethiopia, Krapf quickly immersed himself in language study, learning Oromo (Galla), Amharic, and Ge'ez [Geez]. He translated the Gospels of John and Matthew into Oromo, and wrote his first grammar and dictionary. In Shoa, Krapf became embroiled in a short controversy concerning whether to print Bibles in Amharic or Ge'ez; the British and Foreign Bible Society, which published Krapf's work, argued

General names (of people or things) can serve as predicates, or as subjects in universal sentences. Empty names have no denotation, but can be replaced by synonymous combinations of singular and general terms. An apparent name, Kotarbiński argued, can only legitimately be used if the expression in which it occurs can be translated into an expression containing only genuine names. Thus we can get away with saying *John's anger was shortlived* because we can translate it into *John was not angry for long*. Saying the former no more commits us to the existence of 'anger' than saying *Mary took umbrage at my words* commits us to the existence of 'umbrage'; indeed, all that exist are objects (including people).

Kotarbiński was one of the first philosophers to consider the possibility of a three-valued logic (before Jan Łukasiewicz or Emil Post in the 1920s). He came to this position in 1912 from consideration of the possibility of creative activity. Assuming that truth is eternal, and that a statement S affirming the existence of an object O can be true if and only if O exists, it should be impossible to create O. For until O is created, S cannot be true, therefore – by one version of the law of the excluded middle – S must be (eternally) false and therefore O's existence cannot be affirmed. Kotarbiński rejected this version of the Aristotelian principle, suggesting instead that before O is created, S is neither true nor false, but indefinite.

Bibliography

Kotarbiński T (1965). *Praxiology: an introduction to the science of efficient action.* Warsaw and Oxford: PWN-Pergamon Press.

Kotarbiński T (1966). *Gnosiology: the scientific approach to the theory of knowledge.* Warsaw and Oxford: PWN-Pergamon Press.

Kotarbiński T (1968). 'The problem of the existence of the future.' *The Polish Review 13,* 7–22.

Woleński J (ed.) (1990). *Kotarbiński: logic, semantics, and ontology.* Dordrecht and Boston: Kluwer Academic Publishers.

Koya:ng *See:* Gondi.

Krapf, Johann Ludwig (1810–1881)

S Pugach, Ohio State University, Lima, Ohio

Johann Krapf was born on February 11, 1810, in Derendingen, Württemburg, and died on November 26, 1881, in Korntal, Württemburg. One of four children, he was the son of Johann Jakob Krapf, a farmer, and Katherina Maria Braun. He attended the Osterbergschule in Tübingen, where he discovered a talent for languages, studying Greek, Latin, French, Italian, and Hebrew (Pirouet, 1999). Krapf's next step was training to become a missionary at the Basel Missionary Institute, a training he interrupted to study theology at the University of Tübingen. In 1834 Krapf returned to the mission after a five-year hiatus, in which time he had been ordained and become a parish minister.

At this point the German Krapf joined the British Church Missionary Society, and in 1837 was sent to help found a mission station in Abyssinia. After unsuccessfully trying to establish the station and being driven from the region, Krapf moved to Shoa, in southern Ethiopia, before he was ultimately forced to go even further south, to Mombasa (Miehe, 1983). During his 19 years in Africa, Krapf achieved fame in two arenas: as a writer of historical travelogues, and as a linguist who compiled grammars and biblical translations for several languages. In this second capacity, Krapf belonged to a group of German missionaries, among them S. W. Koelle and J. F. Schön, who introduced the discipline of African linguistics to Europe (*see* **Koelle, Sigismund Wilhelm (1823–1902)**).

Once in Ethiopia, Krapf quickly immersed himself in language study, learning Oromo (Galla), Amharic, and Ge'ez [Geez]. He translated the Gospels of John and Matthew into Oromo, and wrote his first grammar and dictionary. In Shoa, Krapf became embroiled in a short controversy concerning whether to print Bibles in Amharic or Ge'ez; the British and Foreign Bible Society, which published Krapf's work, argued

for Amharic, but Krapf insisted that, because the natives of the area did not like Amharic, biblical translations should instead be printed in both Amharic and Ge'ez, with the languages side by side (Krapf, letter to the British and Foreign Bible Society, Archives of the British and Foreign Bible Society, Cambridge, August 1839). While in Ethiopia, Krapf also gathered samples of Amharic and Ge'ez handwriting, which he sent back to Europe for analysis. In 1842, his achievements in Ethiopia were recognized when his alma mater, the University of Tübingen, awarded him a doctorate for his Ethiopian work.

After Krapf left Ethiopia, he settled in Mombasa, where he lived from 1844 to 1855. Krapf lost his first wife, Rosina, and their daughter to illness during their first year in Kenya. In 1846 he was joined by Johannes Rebmann, who commonly accompanied Krapf on his journeys to the interior, and also collaborated with him on linguistic work. During their journeys, Rebmann and Krapf became the first Europeans to reach Mount Kilimanjaro (Krapf, 1858, 1860). In East Africa, Krapf studied Swahili and several related languages, including Kinika (Kirabai), the language of the Mijikenda. A nine-page biblical translation of Genesis, which Krapf completed not long after his arrival in Mombasa, was the first Swahili text printed in Roman script. In 1850 Krapf compiled a short comparative vocabulary of six East African languages, which also contained ethnographic information on the peoples who spoke them (Krapf, 1850). Krapf commented on the affinity of the six languages, and on their connection to a much larger family of languages spoken throughout sub-Saharan Africa. These were the Bantu languages, though Krapf did not name them as such; his colleague Wilhelm Bleek, Cape Colony governor Sir George Grey's librarian, was instead the first to call the language group Bantu (Frankl, 1992).

Krapf's greatest accomplishments in linguistics are undoubtedly found in his work on Swahili. Krapf was the first to compose a Swahili grammar, which was published in 1850. Frankl suggests that Krapf would have been the first to publish a Swahili dictionary, too, but white ants destroyed much of his research, delaying publication until 1882, a year after Krapf's death. In 1860 Krapf did, however, see a shorter version of his lexicon disseminated by the London Society for the Propagation of the Gospel. Thus, while Bishop Edward Steere's Swahili handbook predated Krapf's complete dictionary (Steere, 1870), Krapf was nonetheless a pioneer in the European study of Swahili (Frankl, 1992).

In 1853 Krapf left Kenya for health reasons, stopping briefly in London before settling permanently in Korntal, where he married twice more before his death in 1881. Krapf maintained his involvement with the mission from afar, advising other missionaries from his home in Germany and making several brief trips to Africa after his retirement. He was in constant communication with the British and Foreign Bible Society, and continued to translate biblical texts into Oromo and Amharic long after he had returned to Germany (Pirouet, 1999).

See also: Koelle, Sigismund Wilhelm (1823–1902); Multi-valued Logics.

Bibliography

Bridges R C (1968). 'Krapf and the strategy of the mission to East Africa.' *Makerere Journal 5,* 7-75.

British and Foreign Bible Society Archive, Cambridge University Library, Correspondence of J. L. Krapf.

Claus W (1882). *Dr. Ludwig Krapf, Weil Missionar in Ostafrika.* Basel: C F Spittler.

Frankl P J L (1992). 'Johann Ludwig Krapf and the birth of Swahili studies.' *Zeitschrift der Deutschen Morgenlandischen Gesellschaft 142(1),* 12–20.

Griefenow-Mewis C (1996). 'J. L. Krapf and his role in researching and describing East African languages.' *Afrikanistische Arbeitspapiere Swahili Forum 3(47).*

Jungraithmayr H & Mohlig W J G (eds.). *Lexikon der Afrikanistik: Afrikanische Sprachen und ihre Erforschung.* Berlin: Dietrich Reimer Verlag.

Krapf J L (1850). *Vocabulary of six East-African languages: Kisuáheli, Kiníka, Kikámba, Kipokómo, Kihiáu, Kigálla.* Tübingen: L. F. Fues.

Krapf J L (1858). *Reisen in Ostafrika: ausgeführt in den Jahren 1837–1855.* Stuttgart: Stroh.

Krapf J L (1860). *Travels, researches, and missionary labours, during an eighteen year's residence in Eastern Africa. Together with journeys to Jagga, Usambara, Ukambani, Shoa, Abessinia, and Khartum; and a coasting voyage from Mombaz to cape Delgado.* London: Trübner.

Miehe G (1983). 'Krapf, Johann Ludwig.' In Jungraithmayr H & Möhlig W J G (eds.) *Lexikon der Afrikanistik: Afrikanische Sprachen und ihre Erforschung.* Berlin: Dietrich Reimer Verlag. 137.

Pirouet M L (1999). 'The legacy of Johann Ludwig Krapf.' *International Bulletin of Missionary Research 23(2).*

Richards C G (1958). *Ludwig Krapf, missionary and explorer.* London: Macmillan.

Steere E (1870). *A handbook of the Swahili language as spoken at Zanzibar.* London: Bell & Daldy.

Kratzenstein, Christian Gottlieb (1723–1795)

C Neis, Potsdam University, Potsdam, Germany

Christian Gottlieb Kratzenstein was born on January 30, 1723 in Wernigerode (Germany) and died on July 7, 1795 in Copenhagen. Kratzenstein was a leading physician, physicist, and natural scientist in the 18th century. Among linguists, he is known as the inventor of a 'speaking machine.'

His father was a school teacher and burgomaster in Wernigerode, where Kratzenstein obtained an excellent education at the local grammar school. In 1742, he started his studies at the University of Halle where he published his first book: *Beweis, dass die Seele ihren Körper baue* in 1743. In this work, he discussed problems like the preformation theory, which regarded organic molecules as determinants of every new creature from the time of Adam and Eve onwards. Kratzenstein discussed the attractive force of Newton in his *Abhandlung von dem Einfluß des Mondes in die Witterungen und in den menschlichen Cörper: Wobey er zugleich zu seinen künftigen Vorlesungen einladet* (1746 and 1747), defended his own *Theoria fluxus diabetici eiusque curandi methodus more geometrice pro gradu Doctoris* (1746) and dealt with worms in the human body (*Abhandlung von der Erzeugung der Würmer im menschlichen Körper*; Kratzenstein, 1748). Most important, however, was the prize essay he sent to the Academy of Bordeaux in 1743 on *Théorie de l'élévation des vapeurs et des exhalaisons, démontrée mathématiquement, qui a remporté le prix, au jugement de l'Académie des belles lettres, sciences et arts.* (Kratzenstein, 1743b), for which he obtained a gold medal. In this treatise he showed that the effluvia (*vapeurs, Dünste*) *which rise from water are invisible air-filled bubbles*. (*These effluvia are invisible as long as being connected in quiet air. But when separated, they become invisible and reflect light according to the laws of Descartes.*) Stimulated by his professor Johann Gottlob Krüger (1715–1759), Kratzenstein wrote his *Schreiben vom Nutzen der Electricität in der Arzneywissenschaft* in 1744 (published in 1772), in which he presented electricity as an omnipotent remedy against every kind of disease. In 1746, he became a doctor of philosophy as well as a doctor of medicine at the University of Halle, where he was subsequently appointed as professor of physics. In 1748, he became a professor of mathematics and mechanics at the Academy of St Petersburg, which attracted many distinguished savants from Western Europe by that time. Kratzenstein was ordered to make mechanical instruments. He conducted experiments with microscopes and optic instruments, supervised patients suffering from diseases in the hospital of the Academy and gave lectures. He also developed new methods of navigation by inventing instruments for the measurement of longitude and latitude degrees and the polar altitude. After long struggles with Mikhail Vasilievich Lomonosov (1714–1765) and other scientists who were opposed to the multitude of foreign professors at the Academy, Kratzenstein left St Petersburg and became a professor of physics in Copenhagen. Until 1786, he gave lectures on different disciplines of the natural sciences, e.g., physics, natural history, botany, zoology, mineralogy, medicine, pharmacology, and chemistry. In 1780, Kratzenstein received the prize of the St Petersburg Academy for his work on artificial human voice. He constructed a 'speech machine' capable of imitating five vowels, which resembled to the *vox humana* register of baroque organs. Kratzenstein was the first to use the free reed and thus inspired the construction of free-reed instruments such as the glass-harmonica, the harmonica, and the accordion.

Kratzenstein was an ingenious scientist whose mechanical reconstruction of the human voice was a first attempt at the creation of artificial voice production.

See also: Lomonosov, Mikhail Vasilyevich (1711–1765).

Bibliography

Engelhardt D von (2002). *Biographische Enzyklopädie deutschsprachiger Mediziner. Bd 1: A-Q.* München: K. G. Saur. 345.

Hirsch A (1886). *Biographisches Lexikon der hervorragenden Aerzte aller Zeiten und Völker.* Wien und Leipzig: Urban & Schwarzenberg. 544.

Killy W & Vierhaus R (eds.) (1997). '*Kratzenstein, Christian Gottlieb, Naturforscher.*' *Deutsche Biographische Enzyklopädie (DBE). Bd 6.* München 73.

Königliche Akademie der Wissenschaften (1969). *Allgemeine Deutsche Biographie.* Siebenzehnter Band. Auf Veranlassung Seiner Majestät des Königs von Bayern herausgegeben durch die historische Commission bei der Königl. Akademie der Wissenschaften. In *Neudruck der 1. Auflage von 1883.* Berlin: Duncker & Humblot. 57–61.

Kratzenstein C G (1743a). *Beweiß, Daß die Seele ihren Cörper baue: In einem Glückwünschungsschreiben an Herrn Heinrich Friedrich Delius; als Derselbe die Doctorwürde in der Arzneygelahrtheit auf der Universität zu Halle annahm.* Halle: Hendel.

Kratzenstein C G (1743b). *Théorie de l'élévation des vapeurs et des exhalaisons, démontrée mathématiquement, qui a remporté le prix, au jugement de l'Académie des belles lettres, sciences et arts.* Bordeaux: Brun.

Kratzenstein Christian G (1745). *Abhandlung von dem Nutzen der Electricität in der Arzneywissenschaft. 2. und vermehrte Aufl.* Halle: Hemmerde.

Kratzenstein C G (1746). 'Theoria fluxus diabetici eiusque curandi methodus more geometrice pro gradu Doctoris.' Halae: Hemmerde. In *Disputationes ad morborum historiam et curationem facientes.* Coll. Alb. Hallerus. Tom. 4.

Kratzenstein C G (1747). *Abhandlung von dem Einfluß des Mondes in die Witterungen und in den menschlichen Cörper : Wobey er zugleich zu seinen künftigen Vorlesungen einladet.* Halle: Hemmerde.

Kratzenstein (1748). *Abhandlung von dem Einfluß des Mondes in die Witherungeu und in den menschlichen Cörper: Wobey er zugleich zu seinen künftigen Vorlesungen einladet. (1746 & 1747).* Halle: Hemmerde.

Kratzenstein C G (1772). *Physicalische Briefe : 1. Von dem Nutzen der Electricität in der Arzneyiwissenschaft 2. Beweis dass die Seele ihren Körper baue.* 4. Auflage. Halle: Hemmerde.

Schott H (ed.) (1997). *Die Chronik der Medizin.* Gütersloh/München: Chronik Verlag.

Snorrason E (1974). *C. G. Kratzenstein. Professor physices experimentalis Petropol. et Havn. and his Studies on Electricity during the Eighteenth Century.* (*Acta historica scientiarum naturalium et medicinalium edidit Bibliotheca universitatis Hauniensis*, vol 29.) Odense: Odense University Press.

Krio

P M K Thomas, University of Oxford, Oxford, UK

The Sociohistorical Development of Krio

Krio, one of the languages spoken in Sierra Leone in West Africa, is a creole language belonging to the Atlantic group of English creoles which are restructured languages with English as the superstrate language and varying degrees of structural influences of the Niger–Congo languages in Africa. It can be further subcategorized as West African on the basis of its areal distribution and immediate linguistic affiliations.

The Domestic and Jamaican hypotheses are two competing views that have been postulated for the origins of Krio. The Domestic Hypothesis, a corollary of monogenesis, is advocated by Hancock (1986) and other creolists. It is argued that Krio is an offshoot of an English variety of creole, which coexisted with and was influenced by a Portuguese-derived pidgin in West Africa. According to Whinnom (1965), this pidgin was related to Sabir, a Romance pidgin and Mediterranean lingua franca spoken between European and non-European sailors and traders from the Middle Ages to the early 20th century. A lesser known version of the Domestic Hypothesis is based on the view stated in E. D. Jones (1956), Berry (1959), and Peterson (1969) that contact between the local inhabitants and various groups of settlers after the founding of the 'Province of Freedom' in the peninsula of Sierra Leone in the 19th century gave rise to Krio.

Sociohistorical developments in the late 18th century and the first half of the 19th century and the attendant linguistic situation around the Sierra Leone peninsula contributed significantly to the establishment and spread of present day Krio. Henry Smeathman, a botanist, who had lived for 3 years on the Banana Island near the Sierra Leone River from 1771 to 1774, proposed the area as suitable for the establishment of an agricultural settlement populated by a free community of equal blacks and whites. Smeathman died before realizing his dream, but his idea was revived by Granville Sharp, one of the philanthropists behind the campaign to repatriate and rehabilitate emancipated African ex-slaves. Following the abolition of slavery in Britain in 1787, emancipated ex-slaves in London became destitute and created a social problem for the British government. Spitzer (1974: 9) describes their destitution and the acute need for repatriation and rehabilitation of these ex-slaves who became known as the Black Poor. A group of philanthropists campaigned vigorously for the identification, purchase, and establishment of a settlement for the Black Poor. Sharp and a group of social reformers known as the The Clapham Sect pioneered the repatriation of 411 Black Poor and some English women to the peninsula of Sierra Leone in 1787.

When they arrived in Sierra Leone, Captain Thompson, who was the leader of the expedition, purchased a piece of land from King Tom, the Temne Chief. The piece of land became known as The Province of Freedom and Thompson called the first settlement Granville Town after Granville Sharp. Historians have attributed the collapse of the first settlement to problems ranging from the fact that the settlers were ill-equipped for the weather to hostilities between the settlers and the Temnes. This occasioned a dispersal of the first settlers so that by 1791 only 48 of the Black Poor remained in The Province of Freedom.

The formation of the Sierra Leone Company and the repatriation of 1131 Africans from Nova Scotia in

Kratzenstein Christian G (1745). *Abhandlung von dem Nutzen der Electricität in der Arzneywissenschaft. 2. und vermehrte Aufl.* Halle: Hemmerde.

Kratzenstein C G (1746). 'Theoria fluxus diabetici eiusque curandi methodus more geometrice pro gradu Doctoris.' Halae: Hemmerde. In *Disputationes ad morborum historiam et curationem facientes.* Coll. Alb. Hallerus. Tom. 4.

Kratzenstein C G (1747). *Abhandlung von dem Einfluß des Mondes in die Witterungen und in den menschlichen Cörper : Wobey er zugleich zu seinen künftigen Vorlesungen einladet.* Halle: Hemmerde.

Kratzenstein (1748). *Abhandlung von dem Einfluß des Mondes in die Witherungeu und in den menschlichen Cörper: Wobey er zugleich zu seinen künftigen Vorlesungen einladet. (1746 & 1747).* Halle: Hemmerde.

Kratzenstein C G (1772). *Physicalische Briefe : 1. Von dem Nutzen der Electricität in der Arzneyiwissenschaft 2. Beweis dass die Seele ihren Körper baue.* 4. Auflage. Halle: Hemmerde.

Schott H (ed.) (1997). *Die Chronik der Medizin.* Gütersloh/ München: Chronik Verlag.

Snorrason E (1974). *C. G. Kratzenstein. Professor physices experimentalis Petropol. et Havn. and his Studies on Electricity during the Eighteenth Century.* (*Acta historica scientiarum naturalium et medicinalium edidit Bibliotheca universitatis Hauniensis*, vol 29.) Odense: Odense University Press.

Krio

P M K Thomas, University of Oxford, Oxford, UK

The Sociohistorical Development of Krio

Krio, one of the languages spoken in Sierra Leone in West Africa, is a creole language belonging to the Atlantic group of English creoles which are restructured languages with English as the superstrate language and varying degrees of structural influences of the Niger–Congo languages in Africa. It can be further subcategorized as West African on the basis of its areal distribution and immediate linguistic affiliations.

The Domestic and Jamaican hypotheses are two competing views that have been postulated for the origins of Krio. The Domestic Hypothesis, a corollary of monogenesis, is advocated by Hancock (1986) and other creolists. It is argued that Krio is an offshoot of an English variety of creole, which coexisted with and was influenced by a Portuguese-derived pidgin in West Africa. According to Whinnom (1965), this pidgin was related to Sabir, a Romance pidgin and Mediterranean lingua franca spoken between European and non-European sailors and traders from the Middle Ages to the early 20th century. A lesser known version of the Domestic Hypothesis is based on the view stated in E. D. Jones (1956), Berry (1959), and Peterson (1969) that contact between the local inhabitants and various groups of settlers after the founding of the 'Province of Freedom' in the peninsula of Sierra Leone in the 19th century gave rise to Krio.

Sociohistorical developments in the late 18th century and the first half of the 19th century and the attendant linguistic situation around the Sierra Leone peninsula contributed significantly to the establishment and spread of present day Krio. Henry Smeathman, a botanist, who had lived for 3 years on the Banana Island near the Sierra Leone River from 1771 to 1774, proposed the area as suitable for the establishment of an agricultural settlement populated by a free community of equal blacks and whites. Smeathman died before realizing his dream, but his idea was revived by Granville Sharp, one of the philanthropists behind the campaign to repatriate and rehabilitate emancipated African ex-slaves. Following the abolition of slavery in Britain in 1787, emancipated ex-slaves in London became destitute and created a social problem for the British government. Spitzer (1974: 9) describes their destitution and the acute need for repatriation and rehabilitation of these ex-slaves who became known as the Black Poor. A group of philanthropists campaigned vigorously for the identification, purchase, and establishment of a settlement for the Black Poor. Sharp and a group of social reformers known as the The Clapham Sect pioneered the repatriation of 411 Black Poor and some English women to the peninsula of Sierra Leone in 1787.

When they arrived in Sierra Leone, Captain Thompson, who was the leader of the expedition, purchased a piece of land from King Tom, the Temne Chief. The piece of land became known as The Province of Freedom and Thompson called the first settlement Granville Town after Granville Sharp. Historians have attributed the collapse of the first settlement to problems ranging from the fact that the settlers were ill-equipped for the weather to hostilities between the settlers and the Temnes. This occasioned a dispersal of the first settlers so that by 1791 only 48 of the Black Poor remained in The Province of Freedom.

The formation of the Sierra Leone Company and the repatriation of 1131 Africans from Nova Scotia in

March 1792 saw the revival of the settlement and the return of some of the first settlers. The Nova Scotians were ex-slaves from the American colonies who had gained their freedom by fighting on the side of the British during the American War of Independence. After the war, they were offered asylum in the British settlement in Nova Scotia. When Thomas Peters, one of the ex-slaves, complained of their ill treatment in Nova Scotia, the British government transported them free to Sierra Leone where, together with the surviving Black Poor, they began a colony they called Free Town. Lieutenant John Clarkson of the Royal Navy, who led the Nova Scotians, became the first Governor of Sierra Leone. Like Granville Town, Free Town (Freetown became the capital of Sierra Leone) also had its fair share of misfortunes. Forty of the Nova Scotians died in the first few weeks of their arrival in the colony. Furthermore, the French razed the settlement to the ground in September 1794, but the surviving settlers rebuilt it.

In 1800, 550 Maroons from Jamaica arrived in the settlement. The Maroons were descendants of slaves who were originally from the Gold Coast but had been taken to Jamaica in the West Indies. They had organized several rebellious campaigns against the British and were promised an amnesty if they surrendered. The British expelled them to Halifax in Nova Scotia but during the bitter winter of 1796–1797, they petitioned to be removed to another place and were taken to Sierra Leone. Instead of being resettled in the Banana Islands, south of the peninsula, they settled in the colony after assisting to foil an uprising in the province. The Maroon population is reported to have dwindled after a decade because some of them died from diseases and others migrated to their original home in the Gold Coast, now known as Ghana.

When the slave trade was prohibited in 1808, Freetown became a Crown Colony. From then on it served as the springboard for British legal and naval operations aimed at combating the slave trade along the West Coast of Africa. Between 1808 and 1864, slave ships were intercepted on the high seas and the redeemed Africans, referred to as Liberated Africans or Recaptives, were released initially in Freetown, and later in other British ports. The Liberated Africans never landed in the New World and, according to Spitzer (1974: 10), "were of heterogeneous ethnic origin, speaking a Babel of African languages." The last batches of settlers were from ethnic tribes in Sierra Leone and other West African countries including Ghana and Nigeria.

The languages spoken in Freetown, including a dialect of West African pidgin/creole, coexisted with English, which was the official language used in the administration of the settlement. Even if historians and creolists disagree on when an English pidgin was first used on the coast of West Africa, or on the ethnolinguistic and demographic composition of the settlement, it is argued here that all of the languages spoken in the settlement, approximately 150 according to Koelle (1854), played roles in the creolization process that produced Krio. The African languages included Temne, the language of the people who sold The Province of Freedom to the first settlers, Mende, Sherbro, Joloff, Bambara, and Kissi. The adstratal influence of settlers from Barbados on the development of Krio probably occurred later, between 1819 and 1896 when, as Berry (1959: 299) points out, convicts from Barbados and disbanded troops from the 2nd and 4th West Indian Regiments were among the early colonists in Sierra Leone from the Caribbean area.

Thus, two periods can be identified in the development of Krio: the pre-1787 period during which a variety of West African creole was spoken in the Sierra Leone estuary and the post-1787 period (between 1787 and the 1860s) during which what is Modern Krio became established out of linguistic input from West African creole, New World creoles, and West African languages. From 1787 onward, the Sierra Leonean variety of West African creole and nascent Krio later converged as a result of pressure from and prestige of the latter. The process, which involved different creole varieties and West African languages, aptly demonstrates the roles of leveling and reconstruction in the creolization model.

Krio spread from Freetown to the interior of Sierra Leone and other West African countries due to the strategic role of Sierra Leone as the base from which Britain spread its colonial administration. The British employed educated Krio and sent them as administrators to other West African colonies in the 19th century. Varieties of Krio are spoken today in the Gambia, Cameroon, Guinea, Senegal, Ghana, Nigeria, and Fernando Po, which is now known as the Island of Bioko.

'Creo' and 'Creole' have been used as variant names for Krio but Fyle and Jones (1980) and Wyse (1980) argue that 'Akiriyo,' a Yoruba term which refers to the Krio habit of paying visits after religious services is the most plausible derivation of Krio. According to Hancock (1969: 19) in the past, 'Creole' was the generic term for the settlers and their descendants, from 1787 to the second half of the 19th century, and their language. At an orthographic conference held in Freetown in April 1984, participants recommended the adoption of Krio as the official designation of the language and name of the people. Krio is today one of the national languages of Sierra

Leone. It is estimated that native speakers of Krio constitute 3% of the population of Sierra Leone and two-thirds of the rest of the population of the country use the language as a lingua franca. The native speakers inhabit mostly Freetown and the western area of the country.

Krio has not decreolized for a variety of reasons. It has been used by writers as a medium of poetry, drama, and short stories. There is a Krio–English dictionary, portions of the Bible including the New Testament have been translated into Krio and many plays performed in Sierra Leone today are written in Krio. Most importantly, Krio has enjoyed an enhanced social status as one of the official languages of Sierra Leone. It is regarded as a full-fledged Sierra Leonean language and not a corrupt bastardized version of English. Along with the other Sierra Leonean languages, it is used in television broadcasts, as well as most business and official public engagements.

Official recognition of the language also extends to education. In the past, an additive bilingual program of language shelter was widely practiced in primary schools throughout the country. Children in primary schools used to be taught to a large extent through the medium of their first language and English was slowly introduced. The program of immersion had the educational aim of enriching the experiences of the children. The indigenous languages and cultures were maintained and further developed as they interacted with English. Sierra Leone's current education policy requires the extension of the use of Krio and the other indigenous languages to secondary education and there is ongoing work to standardize these languages and produce materials for use at this level. This policy will contribute to the extension of the domains of use and the continued stability and spread of Krio as well as the other national languages.

In addition to the other reasons given above, Krio has not decreolized notwithstanding pressure from English, the lexifier language, because it carries no stigma and is an identity marker.

Some Grammatical Features of Krio

Krio uses separate morphemes to express grammatical categories. In general, the plural is marked by the particle *dɛm*, for example:

dɛm pɛn
PLURAL pen
pens
pɛn dɛm
pen PL(URAL)
pens

Whereas English plural forms can be indicated by the suffixes -s, or -es, some Krio words appear to have suffixes attached without necessarily indicating plurality. Such words were acquired from English in their present forms and can be used with reference to both singular and plural objects, for example:

machis
'match'
sus/shuz
'shoe'

The possessive is marked by the particle *in*, for example:

Patrik in buk
Patrick POSSESSIVE book
'Patrick's book'

De or *di* is the progressive marker, for example:

Idit de rait.
Edith PROGRESSIVE write.
'Edith is writing.'

In certain contexts, verbs without preverbal particles express the default past tense but the particle *bin* is the past tense marker, for example:

Banadɛt bin rait.
Bernadette PAST rait.
'Bernadette wrote.'

Go is the future marker, for example:

Lamin go rait.
Lamin FUTURE rait.
'Lamin will write.'

Dɔn is the perfective marker and it combines with *bin* to express the past perfect, for example:

Angela dɔn rait.
Angela PRESENT PERFECT write.
'Angela has written.'
Mamie bin dɔn rait.
Mamie PAST PERFECT write.
'Mamie had written.'

Blan(t) 'used to' is a habitual aspect marker and the modal markers which can cooccur in a sequence as double or multiple modals are *mɔs* 'must,' *fɔ* 'should,' *go* 'intend to, must,' and *kin* 'can, could.' Consider the following examples:

I blant rait.
She HABITUAL ASPECT write.
'She usually writes.'
A bin fɔ mɔs dɔn rait.
I PAST + MODAL + MODAL + PERFECTIVE write.
'I should (would) most certainly have written.'

Some particles are multifunctional and the different functions are determined by context. These particles include:

- *Na* can function as a locative 'in, at, to,' verbal particle 'is,' and adjectival 'that.'
- *ɔt*, a preposition 'out' also functions as verb for 'extinguish, put out.'
- *De* functions as a verb 'to be,' durative marker, and locative adverb 'there.'
- *Bin* functions as a past tense marker and aspect marker.
- *Fɔ* functions as a preposition, modal auxiliary, main verb, and complementizer or infinitive marker.
- *Go* functions as a verb and modal particle.
- *Blant* is not only a past habitual marker 'used to,' but also functions as a main verb 'belong to.'

Other grammatical features of Krio include:

- multiple negation, for example:

  Nɔ tɔk tu am no moh.
  NEGATOR talk to him NEGATOR more.
  'Do not talk to him any more (again).'

- Serial verbs in Krio have the following basic structure:

  NP_1 Aux V_1 (NP_2) V_2 . . .
  Agnɛs bin kuk rɛs gi Josɛf.
  Agnes (NP_1) bin (PAST AUXILIARY) cook (V_1) rice (NP_2) gi (V_2) Joseph.
  'Agnes cooked some rice which she gave to Joseph.'

- focus constructions involving *na* 'it-is,' for example:

  Na Mari bin kɔ'l.
  It-is Mary PAST call.
  'Mary called.'

Some Lexical Features of Krio

Many Krio words are of African origin. The main Sierra Leonean sources of the lexical items include, Mende, Temne, Sherbro, Susu, Yalunka, Limba, Kru, Vai, Fullah, and Mandingo. The other major African sources include, Yoruba, Wolof, Twi, Hausa, and Ibo. Most of the African items have multiple origins and indicate the multiple connections of the language as shown below:

Word	Sources	Meaning
banda	Mende, Kongo, Swahili	a basket made of palmetto straw or of marsh grass and sewn with palmetto, or a thatched house
bara (bala, balanji)	Bambara, Susu, Mandingo	a xylophone
bene	Wolof, Bambara, Mende	benne, sesame
fufu	Twi, Ewe, Wolof, Fon, Mende, and Hausa	mush, wheat flour made into a thin batter and cooked
nyam/ nyamnyam	Wolof, Fullah, Mbundu, Mandingo, Tshiluba, Efik, and Twi	to eat, eat up, food

Other sources of the lexicon of Krio include European languages: mainly English, Portuguese, Spanish, and French. The largest number of words in the lexicon of this creole derives from English. Between 70% and 80% of the Krio lexicon is derived from English and a lot of the words reflect an archaic usage in English. Examples include:

berin 'a funeral, a burial.' Recorded in the OED as *bering(e)*, it is either obsolete or occurs in English dialects.
titi 'girl' and *krabit* 'miserly, mean' are Scottish representations
bre/ribre 'nag' comes from northern English dialect
kɔstamɛnt 'customer' appears in the OED as an obsolete word *customance/custumaunce* in use as far back as 1386, which means 'customary practice; custom, habit, customary gathering, frequenting'
baksay 'buttocks' is a fossil of an earlier English compound *baksyde*, *backeside* meaning 'the hinder or back part, the back, the rear.'

Vulgar words highlight slang and vulgar usages but some of these words have lost their European connotations. Examples include:

pis 'urine'
switpis 'diabetes'
pisbag 'bladder'
pishol 'urethra'

Some of the Krio words represent semantic Africanisms, for instance:

bif 'meat'
fut 'leg, thigh'
met 'co-wife'

Although some Krio words are different from their English etyma, they are related to their etyma via a semantic change through inference, for instance:

bisin 'to care, be concerned about' from 'business'
drap 'arrive unexpectedly' from 'drop'
bot 'to gang up' from 'both'
ton 'penis' from 'stone'

Words derived directly from Portuguese, Spanish, and French include:

pikin 'child' (derived from Portuguese *pequenno* 'little' or Spanish *pequeño*),
bɔku 'plentiful, abundant' (derived from French '*beaucoup*'),

farinha 'flour' (derived from Portuguese but can also be traced to French *farine*)
sabi 'skill, knowledge' (derived from Portuguese *sabir* 'know' or Spanish *saber.* It also occurs in English as *savvy*)
plaba 'quarrel' (derived from Portuguese *palavra* 'words, talk,' Spanish *palabra*, Italian *parola*, French *parole*, and Latin *parabola* 'parable.')
dash 'present' (derived from Portuguese *das-me* 'give me')

Compounds

Some Krio compounds are created by juxtaposition of words of different grammatical categories. Some of the compounds are derived from English phrasal verbs for example:

fɔdɔm 'fall, fall down'
mekes 'hurry'
tayup/tringup 'tie up'

Two or three morpheme parallels in question words include:

wetin 'what'
udat 'who'
usay/wisay 'where'
ustem/wataym 'when'
wetin-du/mek-so 'why'
ɔmɔs/hɔmɔch 'how many'

Other compounds with different bases are instances of the use of metaphoric language through idiomatic calquing, for example:

Adjectives	Body parts (nouns)	Compound
big 'big'	*mɔt* 'mouth'	*bigmɔt* 'boastfulness'
gud 'good'	*bele* 'belly'	*gudbele* 'kind hearted'
big 'big'	*yay* 'eye'	*bigyay* 'greedy'

Other socially and culturally determined compounds include:

santem (sun time) 'midday'
domɔt (door mouth) 'door'
dede-hos (dead house) 'mortuary'
simun 'menstruation'

Some compounds are gender, occupation, actor, and nationality constructions, for example:

umanfɔl 'hen'
bɔy pikin 'male child'
gyal pikin 'female child'
inglishman 'Englishman'
amerikinman 'American'
ganaman 'Ghanaian'

There is also evidence of semantic calques and extensions as the forms borrowed from English gain new meaning. For example:

ɔpstyas 'the brain' (from 'upstairs')
bizi 'menstrual period' (from 'busy')
yad 'toilet' (from 'yard')
big 'older, wealthy, important'

Epenthetic vowels are inserted between consonants where clusters occur in English words, for example: *tikitul* 'kettle'

Reduplication

The following types of complete reduplication have been attested in Krio:

- intensive reduplication of adjectives, adverbs and verbs

Simplex forms	Reduplicated morphemes
tru 'true'	*trutru* 'very true'
dɛ 'there'	*dɛdɛ* 'exactly there; correct'
ay 'high'	*ayay* 'very high'
kwik 'quickly'	*kwikkwik* 'very quickly'

The following reduplicated verbs also indicate an increase in degree and/or intensity:

Simplex forms	Reduplicated morphemes
baŋ 'bang'	*baŋbaŋ* 'very loud noise'
fred 'be frightened'	*fredfred*, 'very frightened'

- iterative/repetitive/frequentative reduplication of verbs

Simplex forms	Reduplicated morphemes
aks 'ask'	*aksaks* 'repeated asking around'
chenj 'change'	*chenjchenj* 'habitually/always/ constantly changing'

- distributive reduplication of numerals

Simplex forms	Reduplicated morphemes
wan 'one'	*wanwan* 'one by one' or 'one to each'
tu 'two'	*tutu* 'two by two' or 'two to each'

- pluralizing reduplication of nouns.

- plurality bordering on uncountability, for example:

Simplex forms	Reduplicated morphemes
af 'half'	*afaf* 'bits', 'halves'
chuk 'thorn'	*chukchuk* 'thorns'

- increased mass, for example:

Simplex form	Reduplicated morpheme
chaf 'chaff'	*chafchaf* 'chaff'

Some Phonological Features of Krio

Consonants

The voiceless dental fricative /θ/ and the voiced dental fricative /ð/ are often occluded. The voiceless dental

fricative /θ/ is reduced to /t/ or in some dialects /f/, for example:

tin 'thin'
tiŋk 'think'

The voiced dental fricative /ð/ is replaced by the alveolar /d/, for example:

da 'that'
dis 'this'
dɛn 'then'

The voiceless glottal fricative /h/ is often omitted in initial positions, for example:

ol 'hold'
os 'house.'

Initial unstressed sounds are omitted in many words, for example:

gri 'agree'
mɛmba 'remember'
chenj 'exchange'
bot/bɒʊt/ 'about'

Consonant clusters involving a fricative and stop in final positions are often reduced, for example:

was was/was 'wasp'
gens 'against'
han 'hand'
fɔs 'first'
brɛs 'breast'

The stop after /l/ is often dropped and the rule appears to be the reduction of the clusters /lt/ and /ld/ to /l/ in final positions for instance:

/*sɔl*/ 'salt'
ol 'old'
wɔl 'world'

The /l/ before a labial consonant or a dental stop at the end of a word is deleted for example:

ɛp/hɛp 'help'
sef/sɛf 'self'

The final voiced stop /d/ in words ending in /nd/ is omitted, for example:

san 'sand'
tan 'stand'
blɛn 'blind'

Like some West African languages such as Mende and Temne, Krio has the following voiceless and voiced labiovelar coarticulated stops:

/kp/
/gb/

Krio has the following prenasalized stops:

/mb/, /nt/, /ŋg/, and /ŋk/

Consonants are palatalized and the different types are:

- nasalized palatal glide

  /ny/

- palatalized alveolars

  /dy/, /sy/, /ty/, and /zy/

- palatalized velars

  /gy/ and /ky/

The sound /v/ is realized as /b/ in certain words, for example:

ib 'heave'
oba 'over'
dreb 'drive'
koba 'cover'

The /v/ in final positions is often rendered as /f/, for example, 'move' is pronounced /muf/.

Adjacent consonants are often transposed. This is known as metathesis and examples include:

/ask/ 'ask' is pronounced /aks/
/rɪsk/ 'risk' is pronounced /riks/
/mɒsk/ 'mosque' is pronounced /mɔks/

Although this is irregular, there is evidence of rhoticism in Krio in a word like /bʁitʃ/ 'bleach' in which /l/ is realized as /r/.

Vowels and Diphthongs

Krio has seven pure vowels and the vowels lack corresponding pairs of long and short vowels. It also has three diphthongs, /ai/, /au/, and /ɔi/. Variants of these diphthongs are /ay/, /aw/, and /oy/. Vowels are introduced to replace English diphthongs for example:

/e/ replaces /eɪ/
/o/ replaces /əʊ/

Vowels are added at the end of a word. These are known as paragogic vowels. Examples include:

gladi 'glad'
dede 'dead'
arata 'rat'
lili 'little'

Suprasegmentals

Krio is a tone language whose intonation is influenced by African tonal languages. There are two tones, a

low tone / ̀/ which is low in all positions and a high tone / ́/. The falling pitch /ˆ/ is used as a realization of the high tone. Tone is syllable-timed and each Krio word or segment has a relevant pitch pattern. Every syllable has a separate tone or relative pitch that is unrelated to stress. The pitch of syllables causes differences in meaning.

See also: Pidgins and Creoles: Overview; Sierra Leone: Language Situation.

Bibliography

Alleyne M C (1980). *Comparative Afro–American: an historical-comparative study of English-based Afro–American dialects of the New World.* Ann Arbor: Karoma.

Berry J (1959). 'The origins of Krio vocabulary.' *Sierra Leone Studies 12,* 298–307.

Bradshaw A T Von (1966). 'A list of Yoruba words in Krio.' *Sierra Leone Language Review 5,* 61–71.

Coomber M E A (1969). *A descriptive study of Krio phonology.* M.S. Thesis. Georgetown University.

Fyle C N & Jones E D (1980). *A Krio–English dictionary.* Oxford: Oxford University Press/Freetown:University Sierra Leone Press.

Greenberg J H (1966). *The languages of Africa.* The Hague: Mouton/Bloomington: Indiana University Press.

Hancock I F (1969). 'A provisional comparison of the English-based Atlantic Creoles.' *African Language Review VIII,* 2–72.

Hancock I F (1986). 'The domestic hypothesis, diffusion and componentiality: an account of Atlantic anglophone creole origins.' In Muysken P & Smith N (eds.) *Substrata versus universals in creole genesis.* Papers from the Amsterdam Creole Workshop, April 1985. Amsterdam/Philadelphia: John Benjamins. 71–102.

Holm J A (2000). *An introduction to pidgins and creoles.* Cambridge, New York: Cambridge University Press.

Huber M & Parkvall M (eds.) (1999). *Spreading the word: the issue of diffusion among the Atlantic creoles.* London: University of Westminster Press.

Hymes D (ed.) (1971). *The pidginization and creolization of languages.* Cambridge: Cambridge University Press.

Jones E D (1956). 'Some aspects of the Sierra Leone patois or Krio.' *Sierra Leone Studies New Series 6,* 97–109.

Jones E D (1968). 'Some tense, mode and aspect markers in Krio.' *African Language Review 7,* 86–89.

Jones E D, Sandred K I *et al.* (eds.) (1992). *Reading and writing Krio: proceedings of a workshop held at the Institute of Public Administration and Management, University of Sierra Leone, Freetown, 29–31 January, 1990.* Uppsala: Almquist & Wiksell International.

Jones F C V (1983). *English-Derived Words in Sierra Leone Krio.* Unpublished Ph.D. thesis, University of Leeds.

Koelle S (1854). *Polyglotta Africana.* Fourah Bay College, University of Sierra Leone. Reprinted 1963.

Kup A P (1961). *A history of Sierra Leone 1400–1787.* Cambridge: Cambridge University Press.

Ladefoged P (1964). *A phonetic study of West African languages.* Cambridge: University Press.

Le Page R B (1960). 'Creole Language Studies I.' London: Macmillan.

Le Page R B (1961). 'Creole Language Studies II.' London: Macmillan.

Mufwene S S (2001). *The ecology of language evolution.* Cambridge, New York: Cambridge University Press.

Peterson J E (1969). *Province of Freetown: a history of Sierra Leone, 1787–1870.* Evanston: Northwestern University Press.

Romaine S (1988). *Pidgin and creole languages.* London: Longman.

Spencer J (ed.) (1971). *The English language in West Africa.* London: Longman.

Spitzer L (1974). *The creoles of Sierra Leone: responses to colonialism 1870–1945.* Madison: University of Wisconsin Press.

Thomas P (2004). *Some syntactic and lexical affinities between Krio and Gullah.* Ph.D. diss., University of Oxford.

Tinelli H (1981). *Creole phonology.* The Hague: Mouton.

Welmers W E (1973). *African language structures.* Berkeley: University of California Press.

Whinnom K (1965). 'The origin of European-based creoles and pidgins.' *Orbis 14,* 509–527.

Wyse A J G (1980). *Search light on the Krio of Sierra Leone: an ethnographical study of a West African people.* Freetown: Institute of African Studies University of Sierra Leone.

Wyse A J G (1989). *The Krio of Sierra Leone: an interpretive history.* London: Hurst in association with the International African Institute.

Kriolu *See:* Cape Verdean Creole.

Kripke, Saul (b. 1940)

F Buekens, Tilburg University, Tilburg, The Netherlands

American logician and philosopher Saul Kripke was born in Omaha, Nebraska in 1940. Kripke is best known for three contributions to philosophy: he developed a semantics for modal logic and published, at an early age, completeness theorems for propositional modal logic (Kripke, 1959) and modal logic with quantification (Kripke, 1963). Reintroducing the venerable Leibnizian notion of a possible world into philosophy, he reestablished analytical metaphysics as a respectable discipline. His 1970 lectures 'Naming and necessity,' later published as a book under the same title in 1980, had an immense influence on contemporary developments in the philosophy of language.

Kripke forcefully argued, with John Stuart Mill and against Gottlob Frege, that ordinary proper names like 'Aristotle' or 'Kurt Gödel' did not refer via a mediating sense or description, but directly. According to the 'picture' Kripke favors – he refused to call his views a theory – a name 'N' refers to a bearer x when the current use of 'N' can be traced back, via a causal–communicative chain, to an initial baptism episode where 'N' is introduced in the language to refer to x. The conceptual apparatus developed in his semantics for modal logic allowed him to reintroduce the distinction between necessity and contingency (a statement is necessarily true only if it is true in all possible words; contingently true if it is true in a world *w* and false in a world *w*'), and to claim, against a wide orthodoxy, that identity statements flanked by names ('rigid designators') are necessarily true, even though they are empirical discoveries: 'Hesperus, the Morning Star, is identical with Phosphorus, the Evening Star,' is a necessary though *a posteriori* truth.

Kripke applied his semantic theories to natural kind terms ('water' or 'tiger'), claiming that the way we fix the reference of natural kind terms may or may not reveal the essential properties of the substances or species they refer to. Kripke's semantic and metaphysical theories provided a framework in which old questions could be reformulated and new questions could be asked: When is a term a natural kind term? What is necessity? What is the relation between conceivability and possibility? Are physical laws metaphysically necessary?

Philosophy of mind was another area in which his ideas gained a foothold (if an essential property of 'pain' is that it 'feels like pain,' can pains be identical with physical states or events)? Kripke also made interesting contributions to the study of the later Wittgenstein in lectures published as *Wittgenstein on rules and private language* (1982), although his work here has been faulted for being not particularly true to the historical Wittgenstein.

See also: Frege, Gottlob (1848–1925); Leibniz, Gottfried Wilhelm (1646–1716); Mill, John Stuart (1806–1873); Modal Logic; Putnam, Hilary (b. 1926).

Bibliography

Kripke S (1959). 'A completeness theorem in modal logic.' *Journal of Symbolic Logic 24(1),* 1–14.

Kripke S (1963). 'Semantical considerations on modal logic.' *Acta Philosophica Fennica 16,* 83–94.

Kripke S (1980). *Naming and necessity.* Cambridge, MA: Harvard.

Kripke S (1977). 'Speaker reference and semantic reference.' In French P, Uehling & Wettstein H (eds.) *Contemporary perspectives in the philosophy of language.* Minneapolis: University of Minnesota Press. 6–27.

Kripke S (1982). *Wittgenstein on rules and private language.* Oxford: Blackwell.

Kroeber, Alfred Louis (1876–1960)

P P Chruszczewski, University of Wrocław, Wrocław, Poland

Alfred Louis Kroeber (Kröber) (1876–1960) was born in Hoboken, New Jersey, on June 11 to upper-middle-class Protestant parents of German ancestry. The family was thoroughly bilingual, with German being spoken at home. During his childhood Kroeber was introduced to Greek and Latin, which made him interested in linguistics. He attended Columbia University, majoring in English (B.A., 1896; M.A., 1897), and took his doctorate there in 1901 under Franz Boas, having presented a 28-page dissertation on 'Decorative symbolism of the Arapacho.'

Kripke, Saul (b. 1940)

F Buekens, Tilburg University, Tilburg, The Netherlands

American logician and philosopher Saul Kripke was born in Omaha, Nebraska in 1940. Kripke is best known for three contributions to philosophy: he developed a semantics for modal logic and published, at an early age, completeness theorems for propositional modal logic (Kripke, 1959) and modal logic with quantification (Kripke, 1963). Reintroducing the venerable Leibnizian notion of a possible world into philosophy, he reestablished analytical metaphysics as a respectable discipline. His 1970 lectures 'Naming and necessity,' later published as a book under the same title in 1980, had an immense influence on contemporary developments in the philosophy of language.

Kripke forcefully argued, with John Stuart Mill and against Gottlob Frege, that ordinary proper names like 'Aristotle' or 'Kurt Gödel' did not refer via a mediating sense or description, but directly. According to the 'picture' Kripke favors – he refused to call his views a theory – a name 'N' refers to a bearer x when the current use of 'N' can be traced back, via a causal–communicative chain, to an initial baptism episode where 'N' is introduced in the language to refer to x. The conceptual apparatus developed in his semantics for modal logic allowed him to reintroduce the distinction between necessity and contingency (a statement is necessarily true only if it is true in all possible words; contingently true if it is true in a world *w* and false in a world *w*'), and to claim, against a wide orthodoxy, that identity statements flanked by names ('rigid designators') are necessarily true, even though they are empirical discoveries: 'Hesperus, the Morning Star, is identical with Phosphorus, the Evening Star,' is a necessary though *a posteriori* truth.

Kripke applied his semantic theories to natural kind terms ('water' or 'tiger'), claiming that the way we fix the reference of natural kind terms may or may not reveal the essential properties of the substances or species they refer to. Kripke's semantic and metaphysical theories provided a framework in which old questions could be reformulated and new questions could be asked: When is a term a natural kind term? What is necessity? What is the relation between conceivability and possibility? Are physical laws metaphysically necessary?

Philosophy of mind was another area in which his ideas gained a foothold (if an essential property of 'pain' is that it 'feels like pain,' can pains be identical with physical states or events)? Kripke also made interesting contributions to the study of the later Wittgenstein in lectures published as *Wittgenstein on rules and private language* (1982), although his work here has been faulted for being not particularly true to the historical Wittgenstein.

See also: Frege, Gottlob (1848–1925); Leibniz, Gottfried Wilhelm (1646–1716); Mill, John Stuart (1806–1873); Modal Logic; Putnam, Hilary (b. 1926).

Bibliography

Kripke S (1959). 'A completeness theorem in modal logic.' *Journal of Symbolic Logic 24(1),* 1–14.

Kripke S (1963). 'Semantical considerations on modal logic.' *Acta Philosophica Fennica 16,* 83–94.

Kripke S (1980). *Naming and necessity.* Cambridge, MA: Harvard.

Kripke S (1977). 'Speaker reference and semantic reference.' In French P, Uehling & Wettstein H (eds.) *Contemporary perspectives in the philosophy of language.* Minneapolis: University of Minnesota Press. 6–27.

Kripke S (1982). *Wittgenstein on rules and private language.* Oxford: Blackwell.

Kroeber, Alfred Louis (1876–1960)

P P Chruszczewski, University of Wrocław, Wrocław, Poland

Alfred Louis Kroeber (Kröber) (1876–1960) was born in Hoboken, New Jersey, on June 11 to upper-middle-class Protestant parents of German ancestry. The family was thoroughly bilingual, with German being spoken at home. During his childhood Kroeber was introduced to Greek and Latin, which made him interested in linguistics. He attended Columbia University, majoring in English (B.A., 1896; M.A., 1897), and took his doctorate there in 1901 under Franz Boas, having presented a 28-page dissertation on 'Decorative symbolism of the Arapacho.'

Kroeber conducted his ethnographic research mainly in the Southwest, the Philippines, the Great Plains, and California, and his archaeological investigations in Mexico and Peru. His linguistic and anthropological studies focused on native Americans (mainly on California Indian languages).

Kroeber was awarded five honorary degrees (Yale, California, Harvard, Columbia, and Chicago). In 1946 he received the Huxley Medal of the Royal Anthropological Society. Kroeber was working on a definition of culture, coming to the conclusion that culture is a product of a historical development (*Anthropology* 1948). He employed a strictly cultural approach and method of analysis while dealing with the history of ancient as well as modern societies. He was of the opinion that, due to numerous parallels between biology and cultural anthropology, cultural phenomena have biological roots ('Historical reconstruction of culture growths and organic evolution' 1931). "The basis of Kroeber's point of view may be called the natural history of culture, with strong emphasis on the humanistic features. That is, Kroeber's main endeavor was to identify the distinguishing features of each culture and to assign them to categories based on the procedures of natural history" (Steward, 1973: 25). Kroeber succeeded in identifying culture as an element of nature also, thanks to his extensive ethnographies of California Indians and their languages (*Handbook of the Indians of California*, 1925) as well as his research on the languages of the South American Indians (*Handbook of South American Indians*, Julian Steward (ed.), 1946–1959).

At the age of 24, Kroeber was offered his first professional position with the California Academy of Sciences in San Francisco. Then in 1906 he was made assistant professor at the University of California at Berkeley (associate professor in 1911 and professor in 1919). During his early years in San Francisco, he underwent several traumatic experiences, including the deaths of his wife and of Ishi, the last survivor of the Yahi-Yana and an excellent source of information concerning Native Americans whom Kroeber had transferred to San Francisco and had supported to live at the Museum. Last but not least of Kroeber's hardships was an ear infection, which resulted in permanent loss of hearing in his left ear.

In 1915–1916, taking his sabbatical year, Kroeber traveled in Europe, coming in contact with Freud's school of analysis in Vienna. Returning to California he became a successful psychoanalyst himself; however, he did not forsake anthropology and gave up psychoanalysis. During his lifetime Kroeber published a total of 460 papers, of which 70 at least partly concern linguistics.

See also: Anthropological Linguistics: Overview; Boas, Franz (1858–1942); Freud, Sigmund (1856–1939).

Bibliography

Count E W (1960). *A. L. Kroeber: a memorial address.* Maxwell Graduate School of Citizenship and Public Affairs: Syracuse University.

Kroeber A L ([1923] 1948). *Anthropology.* New York: Harcourt Brace.

Kroeber A L (1925). *Handbook of the Indians of California.* Smithsonian Institution. *Bulletin. Bureau of American Ethnology*; no. 78. Washington: Government Printing Office.

Kroeber A L (1931). 'Historical reconstruction of culture growths and organic evolution.' *American Anthropologist 33(2),* 149–156.

Kroeber A L (1946). 'The Chibcha.' In Steward J (ed.) *Handbook of South American Indians, vol. 2. The Andean civilizations. Bulletin. Bureau of American Ethnology*; no. 143: 887–909.

Kroeber A L (1952). *On the nature of culture.* Chicago: University of Chicago Press.

Steward J H (1973). *Alfred Kroeber.* New York: Columbia University Press.

Kru Languages

J Bendor-Samuel, High Wycombe, UK

Kru languages are spoken mainly in the forest areas of southwestern Ivory Coast and southern Liberia. Apart from three languages, they form a contiguous block with Kwa languages to their east, Mande languages to their north and west, and Atlantic languages to their west.

Speakers

Reliable population figures are hard to obtain, but it would seem that there are approximately 2 million people who speak one or other of the Kru languages. The three largest language groups are the Guere

Kroeber conducted his ethnographic research mainly in the Southwest, the Philippines, the Great Plains, and California, and his archaeological investigations in Mexico and Peru. His linguistic and anthropological studies focused on native Americans (mainly on California Indian languages).

Kroeber was awarded five honorary degrees (Yale, California, Harvard, Columbia, and Chicago). In 1946 he received the Huxley Medal of the Royal Anthropological Society. Kroeber was working on a definition of culture, coming to the conclusion that culture is a product of a historical development (*Anthropology* 1948). He employed a strictly cultural approach and method of analysis while dealing with the history of ancient as well as modern societies. He was of the opinion that, due to numerous parallels between biology and cultural anthropology, cultural phenomena have biological roots ('Historical reconstruction of culture growths and organic evolution' 1931). "The basis of Kroeber's point of view may be called the natural history of culture, with strong emphasis on the humanistic features. That is, Kroeber's main endeavor was to identify the distinguishing features of each culture and to assign them to categories based on the procedures of natural history" (Steward, 1973: 25). Kroeber succeeded in identifying culture as an element of nature also, thanks to his extensive ethnographies of California Indians and their languages (*Handbook of the Indians of California*, 1925) as well as his research on the languages of the South American Indians (*Handbook of South American Indians*, Julian Steward (ed.), 1946–1959).

At the age of 24, Kroeber was offered his first professional position with the California Academy of Sciences in San Francisco. Then in 1906 he was made assistant professor at the University of California at Berkeley (associate professor in 1911 and professor in 1919). During his early years in San Francisco, he underwent several traumatic experiences, including the deaths of his wife and of Ishi, the last survivor of the Yahi-Yana and an excellent source of information concerning Native Americans whom Kroeber had transferred to San Francisco and had supported to live at the Museum. Last but not least of Kroeber's hardships was an ear infection, which resulted in permanent loss of hearing in his left ear.

In 1915–1916, taking his sabbatical year, Kroeber traveled in Europe, coming in contact with Freud's school of analysis in Vienna. Returning to California he became a successful psychoanalyst himself; however, he did not forsake anthropology and gave up psychoanalysis. During his lifetime Kroeber published a total of 460 papers, of which 70 at least partly concern linguistics.

See also: Anthropological Linguistics: Overview; Boas, Franz (1858–1942); Freud, Sigmund (1856–1939).

Bibliography

Count E W (1960). *A. L. Kroeber: a memorial address.* Maxwell Graduate School of Citizenship and Public Affairs: Syracuse University.

Kroeber A L ([1923] 1948). *Anthropology.* New York: Harcourt Brace.

Kroeber A L (1925). *Handbook of the Indians of California.* Smithsonian Institution. *Bulletin. Bureau of American Ethnology*; no. 78. Washington: Government Printing Office.

Kroeber A L (1931). 'Historical reconstruction of culture growths and organic evolution.' *American Anthropologist 33(2),* 149–156.

Kroeber A L (1946). 'The Chibcha.' In Steward J (ed.) *Handbook of South American Indians, vol. 2. The Andean civilizations. Bulletin. Bureau of American Ethnology*; no. 143: 887–909.

Kroeber A L (1952). *On the nature of culture.* Chicago: University of Chicago Press.

Steward J H (1973). *Alfred Kroeber.* New York: Columbia University Press.

Kru Languages

J Bendor-Samuel, High Wycombe, UK

Kru languages are spoken mainly in the forest areas of southwestern Ivory Coast and southern Liberia. Apart from three languages, they form a contiguous block with Kwa languages to their east, Mande languages to their north and west, and Atlantic languages to their west.

Speakers

Reliable population figures are hard to obtain, but it would seem that there are approximately 2 million people who speak one or other of the Kru languages. The three largest language groups are the Guere

complex (some 400 000 speakers), the Bete complex (approximately 350 000 people), and the Bassa (over 250 000 people).

Kru Studies

Though Sigismund W. Koelle, in 1854, included five Kru languages in his *Polyglotta Africana*, there was little study of these languages until this century. In 1905, Georges Thomann published a grammar of Neyo, but the next substantial work on Kru languages did not appear until 1966, when Gordon Innes published his *Introduction to Grebe* and a Grebe-English dictionary. Subsequently, however, research has begun into many more of the Kru languages, and the work of Marchese (1983, 1986, 1989) has been extensive.

Classification

Greenberg (1963) tentatively included Kru in his Kwa branch of Niger-Congo, but this classification has been rejected by most scholars. Some lexicostatistical studies and the presence of noun class suffixes have suggested that Kru is closer to Gur than to Kwa, though, unlike Gur and Kwa, Kru is nonserializing. More recently, the hypothesis of grouping Kru, Gur, and Adamawa-Ubangi as North Volta-Congo has found increasing acceptance (Williamson and Blench, 2000). However, since systematic lexical study is lacking and in the absence of any conclusive evidence, it seems best to regard Kru as an independent group for the time being.

The Kru languages can be divided into two main groups: eastern and western. Eastern Kru (approximately 500 000 speakers) is spoken exclusively in the Ivory Coast and contains two major subgroups, the Bete complex and the Dida complex. Bete further divides into eastern Bete, spoken in the region of Gagnoa, and western Bete, spoken in a wider area, including Soubre, Guibéroua, Issia, and Daloa. The Dida complex divides into eastern Dida, with six dialects, and western Dida, with two dialect clusters. Kwadia, Bakwe, and Wane are three other eastern Kru languages that do not belong to either the Bete or the Dida complex.

Western Kru (with more than 1 million speakers) is spoken over a considerably larger area compared to eastern Kru and extends from the Sasandra River in the Ivory Coast through southern Liberia. The major division in western Kru is between the Guere complex, Grebo complex, and Bassa. The Guere complex is located in Ivory Coast and comprises some 35 languages/dialects; these may be grouped into four groups: Nyabwa, Wobe, Guere-Krahn, and Konobo. The Grebo complex straddles the Ivory Coast and Liberia border and includes some 25 languages/dialects. In the Ivory Coast, there are two main subgroups: Tepo-Plapo and Pie. In Liberia, there are seven subgroups: Wedebo, Glebo, Jabo, Gedebo, Niaibo, Fopo, and Chedepo. In Liberia, Bassa, with some 15 dialects and some 250 000 speakers, and Klao, with two main dialects, belong to western Kru but lie outside Grebo and Guere.

Three other Kwa languages, Kuwaa, Aizi, and Semɛ, are not grouped with either eastern or western Kru. All three are separate geographically from the rest of the Kru languages. Semɛ is of particular interest since it is located in Burkina Faso, over 300 miles from the nearest Kru language and surrounded by Gur languages. Figure 1, based on work by Marchese, gives a useful summary.

Structural Features

Phonetics and Phonology

Open syllables predominate and words are usually monosyllabic or disyllabic. Vowel sequences occur as do CLV syllables (consonant-L-vowel, where L represents a syllabic l, n, or r). All Kru languages have stops at five points of articulation: bilabial, alveolar, palatal, velar, and labio-velar. A typical Kru vowel system has four front and four back vowels and a central vowel and is marked by vowel harmony with advanced and retracted sets of vowels. Within any one morpheme, only vowels of one set occur.

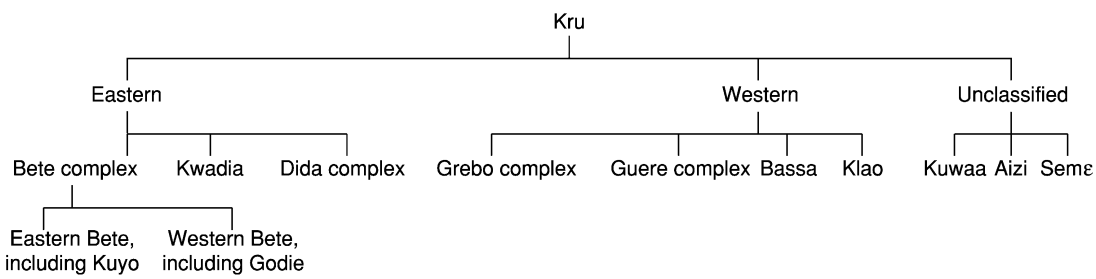


Figure 1 The Kru languages.

Some Kru languages, including Grebo and Krahn, have another type of vowel harmony, with vowels divided into three sets, A, B, and C. Vowels of any one group may cooccur in a morpheme, and vowels of adjacent sets may cooccur. So vowels from set A may occur with vowels from set B, or vowels from set B may occur with vowels from set C, but vowels from set A never cooccur with vowels from set C. Set A consists of i, u, e, and o; set B consists of ɛ, a, and ɔ; and set C consists of *ɩ* and ʋ.

All Kru languages are marked by three or four levels of register tone, which may carry either lexical and/or grammatical functions. Tone, for instance, may distinguish the imperfective from the perfective and the singular from the plural.

Grammar and Syntax

Kru languages have a subject-verb-object (SVO) basic word order, but when an auxiliary is present this changes to S AUX (IO) (DO) V. Kru languages are suffixing. Plurality is often indicated by an *-i* suffix. A number of Kru languages have remnants of noun class suffixes and a few have noun class concord, with agreement between the head and its modifiers. Possessives precede nouns, and body parts function as postpositions indicating direction. Genitives precede nouns, but most modifiers follow their heads.

Kru languages have a basic aspectual system with imperfective and perfective forms marked by suffixes. Progressive and perfect forms are often marked by an auxiliary and so too are conditionals and potential futures. Negation is also frequently signaled by an auxiliary or a particle, though in some languages tone and word order are involved.

See also: Ivory Coast: Language Situation; Kwa Languages; Liberia: Language Situation; Liberia: Language Situation; Mande Languages.

Bibliography

Greenberg J H (1963). *Languages of Africa*. The Hague: Mouton & Co.

Innes G (1966). *Introduction to Grebo*. London: School of Oriental and African Studies, University of London.

Marchese L (1983). *Atlas linguistique kru*. Abidjan: Agence de Coopération Culturelle et Technique, Institut de Linguistique Appliquée, Université d'Abidjan.

Marchese L (1986). *Tense/aspect and the development of auxiliaries in Kru languages*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

Marchese L (1989). 'Kru.' In Bendor-Samuel J (ed.) *The Niger-Congo languages*. Lanham: University Press of America, Inc. 178–215.

Thomann G (1905). *Essai de manuel de la langue néoulé*. Paris.

Williamson K & Blench R (2000). 'Niger-Congo.' In Heine B & Nurse D (eds.) *African languages*. Cambridge: Cambridge University Press. 24–25.

Kruisinga, Etsko (1875–1944)

Roselinde Supheert, Utrecht University, Utrecht, The Netherlands

Etsko Kruisinga was born in Leens, the Netherlands, on December 8, 1875, the son of a merchant. He studied Dutch and English at the University of Groningen, graduating with a *kandidaats* degree (approximately a B.A.) in Dutch philology in 1897, as English at the time had not yet been accepted as an academic degree subject in the Netherlands. After passing additional teacher training examinations (M.O. A and B), Kruisinga went on to teach English and Dutch at a number of secondary schools, until he was pensioned off for health reasons at the age of 48. In 1904, he obtained a doctorate from the University of Bonn for his dissertation on the West Somerset dialect. When the Dutch university statute was changed, enabling students to take degrees in foreign languages, Kruisinga was among the first to obtain a doctorate in English from the University of Amsterdam, in 1922.

Although recognized as a major Anglicist by his contemporaries, Kruisinga was never offered a salaried position at a Dutch university. In addition to his career as a schoolteacher, he worked as an unsalaried lecturer at the University of Utrecht from 1909–1915 and was a driving force in language-teacher training. In 1911, he was one of the founders of the Dutch Association of Modern Language Teachers (*Levende Talen*), and in 1922 he became a lecturer at a private teacher training institution in the Hague, the Modern Language Association (*Vereniging Modern Talen*), later called the School of Linguistic and Literary Studies (*School voor Taal- en Letterkunde*). Here he was appointed principal in 1931 and accepted a controversial professorship in 1936.

Some Kru languages, including Grebo and Krahn, have another type of vowel harmony, with vowels divided into three sets, A, B, and C. Vowels of any one group may cooccur in a morpheme, and vowels of adjacent sets may cooccur. So vowels from set A may occur with vowels from set B, or vowels from set B may occur with vowels from set C, but vowels from set A never cooccur with vowels from set C. Set A consists of i, u, e, and o; set B consists of ɛ, a, and ɔ; and set C consists of ɩ and ʋ.

All Kru languages are marked by three or four levels of register tone, which may carry either lexical and/or grammatical functions. Tone, for instance, may distinguish the imperfective from the perfective and the singular from the plural.

Grammar and Syntax

Kru languages have a subject-verb-object (SVO) basic word order, but when an auxiliary is present this changes to S AUX (IO) (DO) V. Kru languages are suffixing. Plurality is often indicated by an *-i* suffix. A number of Kru languages have remnants of noun class suffixes and a few have noun class concord, with agreement between the head and its modifiers. Possessives precede nouns, and body parts function as postpositions indicating direction. Genitives precede nouns, but most modifiers follow their heads.

Kru languages have a basic aspectual system with imperfective and perfective forms marked by suffixes. Progressive and perfect forms are often marked by an auxiliary and so too are conditionals and potential futures. Negation is also frequently signaled by an auxiliary or a particle, though in some languages tone and word order are involved.

See also: Ivory Coast: Language Situation; Kwa Languages; Liberia: Language Situation; Liberia: Language Situation; Mande Languages.

Bibliography

Greenberg J H (1963). *Languages of Africa*. The Hague: Mouton & Co.

Innes G (1966). *Introduction to Grebo*. London: School of Oriental and African Studies, University of London.

Marchese L (1983). *Atlas linguistique kru*. Abidjan: Agence de Coopération Culturelle et Technique, Institut de Linguistique Appliquée, Université d'Abidjan.

Marchese L (1986). *Tense/aspect and the development of auxiliaries in Kru languages*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.

Marchese L (1989). 'Kru.' In Bendor-Samuel J (ed.) *The Niger-Congo languages*. Lanham: University Press of America, Inc. 178–215.

Thomann G (1905). *Essai de manuel de la langue néoulé*. Paris.

Williamson K & Blench R (2000). 'Niger-Congo.' In Heine B & Nurse D (eds.) *African languages*. Cambridge: Cambridge University Press. 24–25.

Kruisinga, Etsko (1875–1944)

Roselinde Supheert, Utrecht University, Utrecht, The Netherlands

Etsko Kruisinga was born in Leens, the Netherlands, on December 8, 1875, the son of a merchant. He studied Dutch and English at the University of Groningen, graduating with a *kandidaats* degree (approximately a B.A.) in Dutch philology in 1897, as English at the time had not yet been accepted as an academic degree subject in the Netherlands. After passing additional teacher training examinations (M.O. A and B), Kruisinga went on to teach English and Dutch at a number of secondary schools, until he was pensioned off for health reasons at the age of 48. In 1904, he obtained a doctorate from the University of Bonn for his dissertation on the West Somerset dialect. When the Dutch university statute was changed, enabling students to take degrees in foreign languages, Kruisinga was among the first to obtain a doctorate in English from the University of Amsterdam, in 1922.

Although recognized as a major Anglicist by his contemporaries, Kruisinga was never offered a salaried position at a Dutch university. In addition to his career as a schoolteacher, he worked as an unsalaried lecturer at the University of Utrecht from 1909–1915 and was a driving force in language-teacher training. In 1911, he was one of the founders of the Dutch Association of Modern Language Teachers (*Levende Talen*), and in 1922 he became a lecturer at a private teacher training institution in the Hague, the Modern Language Association (*Vereniging Modern Talen*), later called the School of Linguistic and Literary Studies (*School voor Taal- en Letterkunde*). Here he was appointed principal in 1931 and accepted a controversial professorship in 1936.

Throughout his career Kruisinga published an astonishing number of articles, reviews, and books. Together with R. W. Zandvoort, in 1919 he stood at the cradle of the foremost Dutch journal for Anglicists, *English Studies*, which he edited and to which he contributed a series of articles on syntax, among other things. Kruisinga's publications range from phonetics and phonology to English cultural history and editions of literary works. His major contributions to linguistics are in the field of English syntax, notably word groups. Kruisinga's work forms a break with tradition in that he increasingly questioned the relevance of philology at a time when a historical approach was the norm. Thus, for example, in his multivolume *A grammar of present-day English*, which saw numerous later editions as *A handbook of present-day English* and covers syntax, idiom, phonetics, and historical linguistics, the volumes dedicated to syntax draw on authentic sentences from sources such as Thomas Hardy, Elizabeth Gaskell, and the *Times* of London.

In Kruisinga's view, classical languages and the historical study of English did not offer suitable tools for analyzing these samples of a living language. He favored and developed a practical framework, modeled on the language that was to be analyzed, without redundant categories dictated by philology.

Kruisinga's formal approach to grammar reflects the influence of a variety of linguists, such as F. N. Finck, C. C. Uhlenbeck, and H. Sweet, with the latter's distinction between form and meaning being a key influence. Kruisinga's work, in turn, attracted international attention. The later, and more radical, editions of his *Handbook* were widely noticed. The radical nature of his approach, his concentrated style, and his position outside of Dutch university, however, to some extent curtailed his influence.

See also: English, Later Modern (ca. 1700–1900); Poutsma, Hendrik (1856–1937); Sweet, Henry (1845–1912).

Bibliography

Essen A J van (1983). *E. Kruisinga: a chapter in the history of linguistics in the Netherlands.* Leiden: Martinus Nijhoff.

Essen A J van (1991). 'E. Kruisinga.' In Leitner G (ed.) *English traditional grammars.* Amsterdam and Philadelphia: John Benjamins Publishing Company. 153–174.

Kruisinga E (1904). *A grammar of the dialect of West-Somerset.* Bonn: Carl Georgi.

Kruisinga E (1909–1912). *A grammar of present-day English.* Utrecht: Kemink.

Kruisinga E (1917). *An English grammar for Dutch students.* Utrecht: Kemink.

Kruisinga E (1943). *The phonetic structure of English words.* Bern: Francke.

Stuurman F (1993). *English masters and their era.* Amsterdam: Amsterdam University Press.

Supheert R (1995). 'English professionals.' In *Yeats in Holland: the reception of the work of W. B. Yeats in the Netherlands before World War II.* Amsterdam-Atlanta: Rodopi. 24–94.

Kruszewski, Mikołag Habdank (1851–1887)

E F K Koerner, Zentrum für Allgemeine Sprachwissenschaft, Berlin, Germany

Kruszewski was born to Polish parents on December 6, 1851 in Luck (Lutsk), Volhynia, then part of the Ukraine, some 200 miles southeast of Warsaw, and died on November 12, 1887 in Kazan, of a degenerative neurological disease. He went to Warsaw for his university studies, enrolling in psychology and philosophy, and thoroughly absorbed the English empiricist tradition of Bacon (*see* **Bacon, Roger (c. 1214-1292/4)**), Locke (*see* **Locke, John (1632–1704)**), and Hume. After taking his Master's degree with a thesis on incantations in Russian folklore in 1875, he joined Baudouin de Courtenay at the University of Kazan in 1878, having had an intensive correspondence on linguistic issues with him for several years previous to that. At Kazan, Kruszewski soon played an important role in the discussions organized by Baudouin at his home; his background in the behavioral sciences appears to have been one of the reasons for his orientation toward general theory. In his review of Saussure's *Mémoire* (1878; *see* **Saussure, Ferdinand (-Mongin) de (1857–1913)**) and Brugmann's 'Nasalis sonans' article (1876; *see* **Brugmann, Karl (1849–1919)**), he suggested (1880) that a distinction between '*zvuk*' (sound) and '*fonema*' (phoneme) be made, a distinction that made history many years after his death (Kilbury, 1976: 17–23; Klausenburger, 1978). His philosophy of science, however, which was closer to Schleicher (*see* **Schlegel, August Wilhelm von (1767–1845)**) than to the Neogrammarians (notably Hermann Paul; *see* **Paul, Hermann (1846–1921)**), appears to have hampered his success

Throughout his career Kruisinga published an astonishing number of articles, reviews, and books. Together with R. W. Zandvoort, in 1919 he stood at the cradle of the foremost Dutch journal for Anglicists, *English Studies*, which he edited and to which he contributed a series of articles on syntax, among other things. Kruisinga's publications range from phonetics and phonology to English cultural history and editions of literary works. His major contributions to linguistics are in the field of English syntax, notably word groups. Kruisinga's work forms a break with tradition in that he increasingly questioned the relevance of philology at a time when a historical approach was the norm. Thus, for example, in his multivolume *A grammar of present-day English*, which saw numerous later editions as *A handbook of present-day English* and covers syntax, idiom, phonetics, and historical linguistics, the volumes dedicated to syntax draw on authentic sentences from sources such as Thomas Hardy, Elizabeth Gaskell, and the *Times* of London.

In Kruisinga's view, classical languages and the historical study of English did not offer suitable tools for analyzing these samples of a living language. He favored and developed a practical framework, modeled on the language that was to be analyzed, without redundant categories dictated by philology.

Kruisinga's formal approach to grammar reflects the influence of a variety of linguists, such as F. N. Finck, C. C. Uhlenbeck, and H. Sweet, with the latter's distinction between form and meaning being a key influence. Kruisinga's work, in turn, attracted international attention. The later, and more radical, editions of his *Handbook* were widely noticed. The radical nature of his approach, his concentrated style, and his position outside of Dutch university, however, to some extent curtailed his influence.

See also: English, Later Modern (ca. 1700–1900); Poutsma, Hendrik (1856–1937); Sweet, Henry (1845–1912).

Bibliography

Essen A J van (1983). *E. Kruisinga: a chapter in the history of linguistics in the Netherlands.* Leiden: Martinus Nijhoff.

Essen A J van (1991). 'E. Kruisinga.' In Leitner G (ed.) *English traditional grammars.* Amsterdam and Philadelphia: John Benjamins Publishing Company. 153–174.

Kruisinga E (1904). *A grammar of the dialect of West-Somerset.* Bonn: Carl Georgi.

Kruisinga E (1909–1912). *A grammar of present-day English.* Utrecht: Kemink.

Kruisinga E (1917). *An English grammar for Dutch students.* Utrecht: Kemink.

Kruisinga E (1943). *The phonetic structure of English words.* Bern: Francke.

Stuurman F (1993). *English masters and their era.* Amsterdam: Amsterdam University Press.

Supheert R (1995). 'English professionals.' In *Yeats in Holland: the reception of the work of W. B. Yeats in the Netherlands before World War II.* Amsterdam-Atlanta: Rodopi. 24–94.

Kruszewski, Mikotag Habdank (1851–1887)

E F K Koerner, Zentrum für Allgemeine Sprachwissenschaft, Berlin, Germany

Kruszewski was born to Polish parents on December 6, 1851 in Luck (Lutsk), Volhynia, then part of the Ukraine, some 200 miles southeast of Warsaw, and died on November 12, 1887 in Kazan, of a degenerative neurological disease. He went to Warsaw for his university studies, enrolling in psychology and philosophy, and thoroughly absorbed the English empiricist tradition of Bacon (*see* **Bacon, Roger (c. 1214-1292/4)**), Locke (*see* **Locke, John (1632–1704)**), and Hume. After taking his Master's degree with a thesis on incantations in Russian folklore in 1875, he joined Baudouin de Courtenay at the University of Kazan in 1878, having had an intensive correspondence on linguistic issues with him for several years previous to that. At Kazan, Kruszewski soon played an important role in the discussions organized by Baudouin at his home; his background in the behavioral sciences appears to have been one of the reasons for his orientation toward general theory. In his review of Saussure's *Mémoire* (1878; *see* **Saussure, Ferdinand (-Mongin) de (1857–1913)**) and Brugmann's 'Nasalis sonans' article (1876; *see* **Brugmann, Karl (1849–1919)**), he suggested (1880) that a distinction between '*zvuk*' (sound) and '*fonema*' (phoneme) be made, a distinction that made history many years after his death (Kilbury, 1976: 17–23; Klausenburger, 1978). His philosophy of science, however, which was closer to Schleicher (*see* **Schlegel, August Wilhelm von (1767–1845)**) than to the Neogrammarians (notably Hermann Paul; *see* **Paul, Hermann (1846–1921)**), appears to have hampered his success

among the 'mainstreamers' of his day (cf. Koerner, 1986: 380–381, 382–384); it required the reappraisal of his work by Roman Jakobson (1969; *see* **Jakobson, Roman (1896–1982)**) and others to give due recognition to his contribution to the field. Saussure was familiar with Kruszewski's writings, especially those that had appeared in German translation (1881: 1884–1890), and it is there that he could have found a number of ideas that are now familiar from Saussure's *Cours*. These include the arbitrariness of the linguistic sign and the distinction between 'syntagmatic' and 'associative' (since Firth, 1935: 'paradigmatic') relationships, which in Kruszewski's terms were established by the laws of association he had found in Mill's *System of logic* (1843; *see* **Mill, John Stuart (1806–1873)**), namely, those of coexistence and those of similarity. (Only during the 1960s did Jakobson see a parallel between Kruszewski's theory and the two basic types of aphasia.) Kruszewski was also responsible for clarifying the distinction between Lautwandel (sound change) and Lautwechsel (sound alternation), a distinction incorporated by Paul in the second edition of his *Principien* (1886) without due acknowledgment. In other words, Kruszewski was quite clear about the essential difference between the historical and the nonhistorical approach to linguistic phenomena.

Kruszewski's career at the University of Kazan lasted little more than five years – by 1883 when Baudouin had left for Dorpat (now Tartu, Estonia) and he himself had moved up to the highest rung of the academic ladder, his mental illness and physical problems had caught up with him. It was left to Baudouin, his former student Aleksandr I. Aleksandrov, and Friedrich Techmer, editor of the *Internationale Zeitschrift für Allgemeine Sprachwissenschaft*, to complete the German translation of his main work and see it through the press (Kruszewski, 1884–1890). The 1995 translation of his major writings in linguistic theory should ensure recognition of his legacy.

See also: Bacon, Roger (c. 1214-1292/4); Brugmann, Karl (1849–1919); Jakobson, Roman (1896–1982); Locke, John (1632–1704); Mill, John Stuart (1806–1873); Paul, Hermann (1846–1921); Saussure, Ferdinand (-Mongin) de (1857–1913); Schlegel, August Wilhelm von (1767–1845); Ukraine: Language Situation.

Bibliography

Jakobson R (1969). 'L'importanza di Kruszewski per lo sviluppo della linguistica.' *Ricerche slavistiche 13,* 3–23.

Jakobson R (1972). 'The Kazan' school of Polish linguistics and its place in the international development of phonology.' In Jakobson R (ed.) *Selected writings, vol. II: Word and language.* The Hague: Mouton.

Kilbury J (1976). *The development of morphophonemic theory.* Amsterdam: John Benjamins.

Klausenburger J (1978). 'Mikotag Kruszewski's theory of morphophonology.' *HL 5,* 109–120.

Koerner E F K (1986). 'Kruszewski's contribution to general linguistic theory.' In Kastovsky D & Szwedek A (eds.) *Linguistics across historical and geographical boundaries,* vol. 1. Berlin: Mouton de Gruyter.

Kruszewski M (1880). 'Lingvističeskie zametki. Vol. I: Novejšie otkrytija v oblasti ario-evropejskogo vokalizma.' *Russkij filogičeskij vestnik 4,* 33–45.

Kruszewski M (1881). *Über die Lautabwechslung.* Kazan: Universitäts-Buchdruckerei.

Kruszewski M (1884–1890). 'Prinzipien der Sprachentwickelung.' *Internationale Zeitschrift für Allgemeine Sprachwissenschaft 1*, 295–307; *2*, 258–268; *3*, 145–170; *5*, 133–144, 360–399.

Kruszewski M (1995). *Writings in general linguistics: On sound alternation. An outline of linguistic science.* Amsterdam: John Benjamins.

Kuiper, Franciscus Bernardus Jacobus (1907–2003)

J-L Chevillard, CNRS, Université Paris 7, Paris, France

Franciscus Bernardus Jacobus Kuiper was born on July 7, 1907 in the Netherlands. He studied Latin, Greek, and Sanskrit at the University of Leiden. He also took private lessons from the Vedic scholar W. Caland at Utrecht. After completing his dissertation in 1934, Kuiper was high school teacher of Latin and Greek in Indonesia, where he also started to learn Tamil with a native speaker. In 1939, he was appointed professor of Sanskrit at Leiden University, where he taught Sanskrit, Prakrit, and Indo–European linguistics until his retirement in 1972. During World War II, while the university was closed, he deepened his knowledge of Tamil and also studied Munda languages. A recurrent topic in his work is the interaction in a linguistic area between languages belonging to different families, as in his *Proto-Munda words in Sanskrit* (1948) or his 'The Genesis of a Linguistic

among the 'mainstreamers' of his day (cf. Koerner, 1986: 380–381, 382–384); it required the reappraisal of his work by Roman Jakobson (1969; *see* **Jakobson, Roman (1896–1982)**) and others to give due recognition to his contribution to the field. Saussure was familiar with Kruszewski's writings, especially those that had appeared in German translation (1881: 1884–1890), and it is there that he could have found a number of ideas that are now familiar from Saussure's *Cours*. These include the arbitrariness of the linguistic sign and the distinction between 'syntagmatic' and 'associative' (since Firth, 1935: 'paradigmatic') relationships, which in Kruszewski's terms were established by the laws of association he had found in Mill's *System of logic* (1843; *see* **Mill, John Stuart (1806–1873)**), namely, those of coexistence and those of similarity. (Only during the 1960s did Jakobson see a parallel between Kruszewski's theory and the two basic types of aphasia.) Kruszewski was also responsible for clarifying the distinction between Lautwandel (sound change) and Lautwechsel (sound alternation), a distinction incorporated by Paul in the second edition of his *Principien* (1886) without due acknowledgment. In other words, Kruszewski was quite clear about the essential difference between the historical and the nonhistorical approach to linguistic phenomena.

Kruszewski's career at the University of Kazan lasted little more than five years – by 1883 when Baudouin had left for Dorpat (now Tartu, Estonia) and he himself had moved up to the highest rung of the academic ladder, his mental illness and physical problems had caught up with him. It was left to Baudouin, his former student Aleksandr I. Aleksandrov, and Friedrich Techmer, editor of the *Internationale Zeitschrift für Allgemeine Sprachwissenschaft*, to complete the German translation of his main work and see it through the press (Kruszewski, 1884–1890). The 1995 translation of his major writings in linguistic theory should ensure recognition of his legacy.

See also: Bacon, Roger (c. 1214-1292/4); Brugmann, Karl (1849–1919); Jakobson, Roman (1896–1982); Locke, John (1632–1704); Mill, John Stuart (1806–1873); Paul, Hermann (1846–1921); Saussure, Ferdinand (-Mongin) de (1857–1913); Schlegel, August Wilhelm von (1767–1845); Ukraine: Language Situation.

Bibliography

Jakobson R (1969). 'L'importanza di Kruszewski per lo sviluppo della linguistica.' *Ricerche slavistiche 13,* 3–23.

Jakobson R (1972). 'The Kazan' school of Polish linguistics and its place in the international development of phonology.' In Jakobson R (ed.) *Selected writings, vol. II: Word and language.* The Hague: Mouton.

Kilbury J (1976). *The development of morphophonemic theory.* Amsterdam: John Benjamins.

Klausenburger J (1978). 'Mikotag Kruszewski's theory of morphophonology.' *HL 5,* 109–120.

Koerner E F K (1986). 'Kruszewski's contribution to general linguistic theory.' In Kastovsky D & Szwedek A (eds.) *Linguistics across historical and geographical boundaries,* vol. 1. Berlin: Mouton de Gruyter.

Kruszewski M (1880). 'Lingvističeskie zametki. Vol. I: Novejšie otkrytija v oblasti ario-evropejskogo vokalizma.' *Russkij filogičeskij vestnik 4,* 33–45.

Kruszewski M (1881). *Über die Lautabwechslung.* Kazan: Universitäts-Buchdruckerei.

Kruszewski M (1884–1890). 'Prinzipien der Sprachentwickelung.' *Internationale Zeitschrift für Allgemeine Sprachwissenschaft 1*, 295–307; *2*, 258–268; *3*, 145–170; *5*, 133–144, 360–399.

Kruszewski M (1995). *Writings in general linguistics: On sound alternation. An outline of linguistic science.* Amsterdam: John Benjamins.

Kuiper, Franciscus Bernardus Jacobus (1907–2003)

J-L Chevillard, CNRS, Université Paris 7, Paris, France

Franciscus Bernardus Jacobus Kuiper was born on July 7, 1907 in the Netherlands. He studied Latin, Greek, and Sanskrit at the University of Leiden. He also took private lessons from the Vedic scholar W. Caland at Utrecht. After completing his dissertation in 1934, Kuiper was high school teacher of Latin and Greek in Indonesia, where he also started to learn Tamil with a native speaker. In 1939, he was appointed professor of Sanskrit at Leiden University, where he taught Sanskrit, Prakrit, and Indo–European linguistics until his retirement in 1972. During World War II, while the university was closed, he deepened his knowledge of Tamil and also studied Munda languages. A recurrent topic in his work is the interaction in a linguistic area between languages belonging to different families, as in his *Proto-Munda words in Sanskrit* (1948) or his 'The Genesis of a Linguistic

Area' (1967), further discussed in Hock (1982), in which he proposes that the Sanskrit quotative, marked by *iti*, owes its origin to Dravidian influence. He also worked on mythology (see Kuiper, 1983) and on theater (see Kuiper, 1979). The beginning of his very extensive bibliography is available in his 1968 Festschrift, edited by Heesterman *et al.*, and continued in several issues of the *Indo–Iranian Journal* (IIJ 19, IIJ 30, IIJ 40), which was founded in 1957 by Kuiper, together with J. W. de Jong. Kuiper was also one of the vice presidents of the International Association of Tamil Research. A detailed account of his work and his career by one of his disciples is available online (see Witzel, 2004).

See also: Bloch, Jules (1880–1953); Burrow, Thomas (1909–1986); Dravidian Languages; Emeneau, Murray Barnson (b. 1904); Sanskrit.

Bibliography

Heesterman J C, Schokker G H & Subramoniam V I (eds.) (1968). *Pratidānam: Indian, Iranian, and Indo–European studies presented to Franciscus Bernardus Jacobus Kuiper on his sixtieth birthday.* The Hague: Mouton de Gruyter, 1968.

Hock H H (1982). 'The Sanskrit Quotative: A Historical and Comparative Study.' *Studies in Linguistic Sciences 12,* 39–85.

Kuiper F B J (1942). *Notes on Vedic noun-inflexion, Mededeelingen der Nederlandsche Akademie van Wetenschappen, Afdeeling Letterkunde.* Nieuwe Reeks; deel 5, no. 4. Amsterdam: Noord-Hollandsche uitgevers maatschappij.

Kuiper F B J (1948). *Proto-Munda words in Sanskrit, Verhandeling der Koninklijke Nederlandse Akademie van Wetenschappen, Afdeeling Letterkunde.* Nieuwe reeks, deel 51, no. 3. Amsterdam: Noord-Hollandsche uitgevers maatschappij.

Kuiper F B J (1962). *Nahali: a comparative study, Mededeelingen der Koninklijke Nederlandse Akademie van Wetenschappen, Afdeeling Letterkunde.* Nieuwe reeks, deel 25, no. 5. Amsterdam: Noord-Hollandsche uitgevers maatschappij.

Kuiper F B J (1967). 'The Genesis of a Linguistic Area.' *Indo–Iranian Journal 10,* 81–102.

Kuiper F B J (1978). *On Zarathustra's language.* Amsterdam and New York: Noord-Hollandsche UM.

Kuiper F B J (1979). *Varuna and Vidūṣaka: on the origin of the Sanskrit drama, Verhandelingen der Koninklijke Nederlandse Akademie van Wetenschappen, Afdeeling Letterkunde.* Nieuwe reeks, deel 10. Amsterdam: North-Holland.

Kuiper F B J (1983). *Ancient Indian cosmogony; essays selected and introduced by John Irwin.* New Delhi: Vikas.

Kuiper F B J (1991). *Aryans in the Rigveda, Leiden studies in Indo-European 1.* Amsterdam: Rodopi.

Kuiper F B J (1997). *Selected writings on Indian linguistics and philology, Leiden studies in Indo–European 8.* Amsterdam: Rodopi.

Witzel M (2004). 'F. B. J. Kuiper (1907–2003).' *Electronic Journal of Vedic Studies 11–1,* 1–17. http://users.primushost.com/~india/ejvs/ejvs1101/ejvs1101Kuiper.pdf.

Kulanapan *See:* Pomoan Languages.

Külpe, Oswald (1862–1915)

D Samain, Université Paris 7, Paris, France

Oswald Külpe was a German philosopher and psychologist who was born in Candau (Latvia) on March 9, 1862, and died in Munich on December 30, 1915. After studying history in Leipzig (1881), where he took an interest in philosophy and psychology, Külpe wandered to Berlin (1882–1884), Göttingen (1884–1885), and Dorpat (nowadays Tartu, Estonia) (1885–1886) before finally returning in 1886 to Wundt's seminar in Leipzig, where he became Wundt's assistant. After his dissertation (1887) and his *Habilitation* (1888), Külpe was made a Privatdozent with teaching responsibilities in psychology. In 1894, he was made the chair of philosophy and aesthetics at the University of Würzburg, where he founded the Institute of Psychology together with Karl Marbe (1896). His career in Würzburg lasted until 1909, when he became professor of philosophy in Bonn. He finally left for Munich in 1913, where he died of influenza contracted while visiting wounded German soldiers.

Area' (1967), further discussed in Hock (1982), in which he proposes that the Sanskrit quotative, marked by *iti*, owes its origin to Dravidian influence. He also worked on mythology (see Kuiper, 1983) and on theater (see Kuiper, 1979). The beginning of his very extensive bibliography is available in his 1968 Festschrift, edited by Heesterman *et al.*, and continued in several issues of the *Indo–Iranian Journal* (IIJ 19, IIJ 30, IIJ 40), which was founded in 1957 by Kuiper, together with J. W. de Jong. Kuiper was also one of the vice presidents of the International Association of Tamil Research. A detailed account of his work and his career by one of his disciples is available online (see Witzel, 2004).

See also: Bloch, Jules (1880–1953); Burrow, Thomas (1909–1986); Dravidian Languages; Emeneau, Murray Barnson (b. 1904); Sanskrit.

Bibliography

Heesterman J C, Schokker G H & Subramoniam V I (eds.) (1968). *Pratidānam: Indian, Iranian, and Indo–European studies presented to Franciscus Bernardus Jacobus Kuiper on his sixtieth birthday.* The Hague: Mouton de Gruyter, 1968.

Hock H H (1982). 'The Sanskrit Quotative: A Historical and Comparative Study.' *Studies in Linguistic Sciences 12,* 39–85.

Kuiper F B J (1942). *Notes on Vedic noun-inflexion, Mededeelingen der Nederlandsche Akademie van Wetenschappen, Afdeeling Letterkunde.* Nieuwe Reeks; deel 5, no. 4. Amsterdam: Noord-Hollandsche uitgevers maatschappij.

Kuiper F B J (1948). *Proto-Munda words in Sanskrit, Verhandeling der Koninklijke Nederlandse Akademie van Wetenschappen, Afdeeling Letterkunde.* Nieuwe reeks, deel 51, no. 3. Amsterdam: Noord-Hollandsche uitgevers maatschappij.

Kuiper F B J (1962). *Nahali: a comparative study, Mededeelingen der Koninklijke Nederlandse Akademie van Wetenschappen, Afdeeling Letterkunde.* Nieuwe reeks, deel 25, no. 5. Amsterdam: Noord-Hollandsche uitgevers maatschappij.

Kuiper F B J (1967). 'The Genesis of a Linguistic Area.' *Indo–Iranian Journal 10,* 81–102.

Kuiper F B J (1978). *On Zarathustra's language.* Amsterdam and New York: Noord-Hollandsche UM.

Kuiper F B J (1979). *Varuna and Vidūṣaka: on the origin of the Sanskrit drama, Verhandelingen der Koninklijke Nederlandse Akademie van Wetenschappen, Afdeeling Letterkunde.* Nieuwe reeks, deel 10. Amsterdam: North-Holland.

Kuiper F B J (1983). *Ancient Indian cosmogony; essays selected and introduced by John Irwin.* New Delhi: Vikas.

Kuiper F B J (1991). *Aryans in the Rigveda, Leiden studies in Indo-European 1.* Amsterdam: Rodopi.

Kuiper F B J (1997). *Selected writings on Indian linguistics and philology, Leiden studies in Indo–European 8.* Amsterdam: Rodopi.

Witzel M (2004). 'F. B. J. Kuiper (1907–2003).' *Electronic Journal of Vedic Studies 11–1,* 1–17. http://users.primushost.com/~india/ejvs/ejvs1101/ejvs1101Kuiper.pdf.

Kulanapan *See:* Pomoan Languages.

Külpe, Oswald (1862–1915)

D Samain, Université Paris 7, Paris, France

Oswald Külpe was a German philosopher and psychologist who was born in Candau (Latvia) on March 9, 1862, and died in Munich on December 30, 1915. After studying history in Leipzig (1881), where he took an interest in philosophy and psychology, Külpe wandered to Berlin (1882–1884), Göttingen (1884–1885), and Dorpat (nowadays Tartu, Estonia) (1885–1886) before finally returning in 1886 to Wundt's seminar in Leipzig, where he became Wundt's assistant. After his dissertation (1887) and his *Habilitation* (1888), Külpe was made a Privatdozent with teaching responsibilities in psychology. In 1894, he was made the chair of philosophy and aesthetics at the University of Würzburg, where he founded the Institute of Psychology together with Karl Marbe (1896). His career in Würzburg lasted until 1909, when he became professor of philosophy in Bonn. He finally left for Munich in 1913, where he died of influenza contracted while visiting wounded German soldiers.

Though Külpe also published on epistemology, logic, and ethics, he is best known as the leading figure of the Würzburg School. He had many prominent researchers among his students, notably Karl Bühler. Influenced by Machian positivism, his *Outlines of psychology* (1893) still classified psychology among natural sciences, proposing to widen experimental methods to upper mental processes (judgment, knowledge, understanding), and thus rejecting some of Wundt's main principles (such as sharing between 'immediate' and 'mediate' experience, the concept of 'psychic causality,' and the strict limitation of the experimental field).

The break-off remained on other grounds after 1896, when Külpe began his experiments on 'systematic self-observation,' according to which subjects had to solve cognitive problems but were also asked to describe the psychic processes happening within themselves. This departure from the careful experimental control over introspection advocated by Wundt was first supposed to reveal 'elements' of thinking similar to those of sensation and feeling. However, as no such elements were found, it finally brought about the inverse conclusions that thought processes could not be explained as resulting from associative connections between images alone, that objective meanings could be experienced without being embodied by specific words or symbols, and thus that imageless factors (e.g., in the form of rules) could therefore determine thought processes. Conversely, Külpe's concept of *Bewusstheit* ('awareness') emphasizes the fact that the meanings of abstract words can be discovered in consciousness even when only these words themselves are perceivable entities. Other experiments stressed the influence of predispositions to respond to stimuli in such and such manner. Külpe drew the conclusion that the process of abstraction depends not only on the material presented to sensation but also on apprehension by the subject, hence abandoning the sensationalist psychology of contents for a psychology recognizing both contents (*Gedanken*) and acts of mind (*Denken*), close to Husserl and Brentano.

See also: Bühler, Karl (1879–1963); Husserl, Edmund (1859–1938); Wundt, Wilhelm (1832–1920).

Bibliography

Danziger K (1980). 'The History of introspection reconsidered.' *Journal of the History of the Behavioral Sciences 16,* 241–262.

Gundlach H (1999). 'Oswald Külpe und die Würzburger Schule.' In Janke W & Schneider W (eds.) *Hundert Jahre Institut für Psychologie und Würzburger Schule.* Göttingen: Hogrefe. 107–124.

Hammer S (1994). *Denkpsychologie – kritischer Realismus. Eine wissenschaftshistorische Studie zum Werk Oswald Külpes.* Frankfurt/Main: Peter Lang Europäischer Verlag der Wissenschaften.

Külpe O (1888). *Die Lehre vom Willen in der modern Psychologie.* Leipzig: Engelmann.

Külpe O (1893). *Grundriss der Psychologie auf experimenteller Grundlage dargestellt.* Leipzig: Engelmann.

Külpe O (1895). *Einleitung in die Philosophie.* Leipzig, S. Hirzel.

Külpe O (1912, 1920, 1921). *Die Realisierung, ein Beitrag zu Grundlegung der Realwissenschaften I, II, and III.* Leipzig: August Messer.

Külpe O (1920). *Vorlesungen über Psychologie* (ed. K Bühler). Leipzig: Hirzel.

Kumyk *See:* Turkic Languages.

Kunihiro, Tetsuya (b. 1929)

Y Nagashima, Gakushuin University, Tokyo, Japan

Tetsuya Kunihiro was born in the city of Ube, Yamaguchi Prefecture, Japan, on July 23, 1929. He graduated with an M.A. in 1954 from the University of Tokyo and received his D.Litt. in 1979 from the same university. After teaching English at Hagi Senior High School in Yamaguchi Prefecture from 1954–1959, he was appointed lecturer in English linguistics at the University of Shimane in 1959 and associate professor in 1964. In 1967 he moved to the University of Ibaraki and in 1968 to the department

Though Külpe also published on epistemology, logic, and ethics, he is best known as the leading figure of the Würzburg School. He had many prominent researchers among his students, notably Karl Bühler. Influenced by Machian positivism, his *Outlines of psychology* (1893) still classified psychology among natural sciences, proposing to widen experimental methods to upper mental processes (judgment, knowledge, understanding), and thus rejecting some of Wundt's main principles (such as sharing between 'immediate' and 'mediate' experience, the concept of 'psychic causality,' and the strict limitation of the experimental field).

The break-off remained on other grounds after 1896, when Külpe began his experiments on 'systematic self-observation,' according to which subjects had to solve cognitive problems but were also asked to describe the psychic processes happening within themselves. This departure from the careful experimental control over introspection advocated by Wundt was first supposed to reveal 'elements' of thinking similar to those of sensation and feeling. However, as no such elements were found, it finally brought about the inverse conclusions that thought processes could not be explained as resulting from associative connections between images alone, that objective meanings could be experienced without being embodied by specific words or symbols, and thus that imageless factors (e.g., in the form of rules) could therefore determine thought processes. Conversely, Külpe's concept of *Bewusstheit* ('awareness') emphasizes the fact that the meanings of abstract words can be discovered in consciousness even when only these words themselves are perceivable entities. Other experiments stressed the influence of predispositions to respond to stimuli in such and such manner. Külpe drew the conclusion that the process of abstraction depends not only on the material presented to sensation but also on apprehension by the subject, hence abandoning the sensationalist psychology of contents for a psychology recognizing both contents (*Gedanken*) and acts of mind (*Denken*), close to Husserl and Brentano.

See also: Bühler, Karl (1879–1963); Husserl, Edmund (1859–1938); Wundt, Wilhelm (1832–1920).

Bibliography

Danziger K (1980). 'The History of introspection reconsidered.' *Journal of the History of the Behavioral Sciences 16,* 241–262.

Gundlach H (1999). 'Oswald Külpe und die Würzburger Schule.' In Janke W & Schneider W (eds.) *Hundert Jahre Institut für Psychologie und Würzburger Schule*. Göttingen: Hogrefe. 107–124.

Hammer S (1994). *Denkpsychologie – kritischer Realismus. Eine wissenschaftshistorische Studie zum Werk Oswald Külpes*. Frankfurt/Main: Peter Lang Europäischer Verlag der Wissenschaften.

Külpe O (1888). *Die Lehre vom Willen in der modern Psychologie*. Leipzig: Engelmann.

Külpe O (1893). *Grundriss der Psychologie auf experimenteller Grundlage dargestellt*. Leipzig: Engelmann.

Külpe O (1895). *Einleitung in die Philosophie*. Leipzig, S. Hirzel.

Külpe O (1912, 1920, 1921). *Die Realisierung, ein Beitrag zu Grundlegung der Realwissenschaften I, II, and III*. Leipzig: August Messer.

Külpe O (1920). *Vorlesungen über Psychologie* (ed. K Bühler). Leipzig: Hirzel.

Kumyk *See:* Turkic Languages.

Kunihiro, Tetsuya (b. 1929)

Y Nagashima, Gakushuin University, Tokyo, Japan

Tetsuya Kunihiro was born in the city of Ube, Yamaguchi Prefecture, Japan, on July 23, 1929. He graduated with an M.A. in 1954 from the University of Tokyo and received his D.Litt. in 1979 from the same university. After teaching English at Hagi Senior High School in Yamaguchi Prefecture from 1954–1959, he was appointed lecturer in English linguistics at the University of Shimane in 1959 and associate professor in 1964. In 1967 he moved to the University of Ibaraki and in 1968 to the department

of linguistics at the University of Tokyo, where he became a professor in 1980 and remained until his retirement in 1990 with the title of Emeritus Professor. During this time he became chairman of the department and served as president of the Linguistic Society of Japan (1985–1988). He has been director of the Tokyo Institute for the Advanced Studies of Language since 1995.

His main field of work is descriptive semantics, especially in the field of lexical and grammatical analysis of Japanese and English. His work on the contrastive study of Japanese and English (Kunihiro, 1967) was pioneering in this field, and he has edited five volumes on the same theme (Kunihiro, 1980). He has also compiled an excellent English–Japanese dictionary as chief editor. As a perceptive analyst of word meaning, he has produced with his colleagues, three volumes on the meaning of basic Japanese words.

By developing the theory of sememics, originated by his teacher Shiro Hattori (*see* **Hattori, Shiro (1908–1995)**), he has established his own theory and descriptive method of lexical meaning with an emphasis on the cognitive aspect and the situational context of the speech event.

See also: Hattori, Shiro (1908–1995); Japan: History of Linguistics; Japanese; Lexical Semantics: Overview.

Bibliography

Kunihiro T (1967). *Structural semantics – contrastive study of Japanese and English*. Sanseido: Tokyo.

Kunihiro T (1970). *Aspects of meaning*. Sanseido: Tokyo.

Kunihiro T (ed.) (1980–1982). *Contrastive studies of Japanese and English* (5 vols). Taishukan Shoten: Tokyo.

Kunihiro T (1997). *An ideal dictionary of Japanese language*. Taishukan Shoten: Tokyo.

Tsuchida S, Uwano Z & Yamada S (eds.) (1990). *Between grammar and meaning. Festschrift for Professor Tetsuya Kunihiro on the occasion of his sixtieth birthday.* Kuroshio Shuppan: Tokyo.

Yamada S, Kikuchi Y & Momiyama Y (eds.) (2000). *Japanese – landscape of meaning and grammar, Festschrift for Professor Tetsuya Kunihiro on the occasion of his seventieth birthday.* Tokyo: Hitsuji Shobo.

Kurath, Hans (1891–1992)

M Pierce, University of Michigan, Ann Arbor, MI, USA

Hans Kurath was born on December 13, 1891, in Villach, Austria, and died in Ann Arbor, Michigan, on January 2, 1992. Kurath came to the United States in 1907, becoming an American citizen in 1912. He was educated at the German-American Teachers Seminar in Milwaukee, Wisconsin, and the Universities of Wisconsin, Texas (A.B. 1914), and Chicago (Ph.D. 1920). Kurath taught German at the University of Texas from 1914 to 1916, and from 1917 to 1919, and taught from 1920 to 1927 at Northwestern University. In 1927 he became professor of German and linguistics at Ohio State University, and in 1930 he was named director of the Linguistic Atlas of the United States. He left Ohio State for Brown University in 1931, where he served as professor of Germanics and linguistics, as well as chair of the department of German and the division of modern languages. In 1946, he became professor of English and editor of the Middle English Dictionary (hereafter MED) at the University of Michigan. Kurath was president of the Linguistic Society of America in 1942 and directed the society's summer Linguistic Institutes from 1947 to 1950. He was the recipient of various scholarly honors; he received honorary doctorates from the University of Chicago and the University of Wisconsin at Milwaukee, as well as the University of Michigan Press Book Award in 1972 for the MED (which he shared with his successors as editor). The esteem his students and colleagues held him in is evident from the volume of papers dedicated to him (Scholler and Reidy, 1973). Kurath retired from Michigan in 1962 but remained scholarly active until his death.

Kurath had a major effect on two fields of linguistics: dialectology and lexicography. His dialect work was rigorously based on Neogrammarian ideas about sound change and resembled similar European projects in issues like choice of linguistic consultants. In addition to directing the Linguistic Atlas project, Kurath published numerous groundbreaking works in this area (e.g., Kurath [1972] and Kurath and McDavid [1961]), a number of his fieldworkers later became prominent dialectologists in their own rights, and the data collected under his direction continue to be useful to dialectologists and sociolinguists. His lexicographical work is also substantial. Ironically, given the overwhelmingly positive reception of the MED, there was some resistance to his

of linguistics at the University of Tokyo, where he became a professor in 1980 and remained until his retirement in 1990 with the title of Emeritus Professor. During this time he became chairman of the department and served as president of the Linguistic Society of Japan (1985–1988). He has been director of the Tokyo Institute for the Advanced Studies of Language since 1995.

His main field of work is descriptive semantics, especially in the field of lexical and grammatical analysis of Japanese and English. His work on the contrastive study of Japanese and English (Kunihiro, 1967) was pioneering in this field, and he has edited five volumes on the same theme (Kunihiro, 1980). He has also compiled an excellent English–Japanese dictionary as chief editor. As a perceptive analyst of word meaning, he has produced with his colleagues, three volumes on the meaning of basic Japanese words.

By developing the theory of sememics, originated by his teacher Shiro Hattori (*see* **Hattori, Shiro (1908–1995)**), he has established his own theory and descriptive method of lexical meaning with an emphasis on the cognitive aspect and the situational context of the speech event.

See also: Hattori, Shiro (1908–1995); Japan: History of Linguistics; Japanese; Lexical Semantics: Overview.

Bibliography

Kunihiro T (1967). *Structural semantics – contrastive study of Japanese and English*. Sanseido: Tokyo.

Kunihiro T (1970). *Aspects of meaning*. Sanseido: Tokyo.

Kunihiro T (ed.) (1980–1982). *Contrastive studies of Japanese and English* (5 vols). Taishukan Shoten: Tokyo.

Kunihiro T (1997). *An ideal dictionary of Japanese language*. Taishukan Shoten: Tokyo.

Tsuchida S, Uwano Z & Yamada S (eds.) (1990). *Between grammar and meaning. Festschrift for Professor Tetsuya Kunihiro on the occasion of his sixtieth birthday.* Kuroshio Shuppan: Tokyo.

Yamada S, Kikuchi Y & Momiyama Y (eds.) (2000). *Japanese – landscape of meaning and grammar, Festschrift for Professor Tetsuya Kunihiro on the occasion of his seventieth birthday.* Tokyo: Hitsuji Shobo.

Kurath, Hans (1891–1992)

M Pierce, University of Michigan, Ann Arbor, MI, USA

Hans Kurath was born on December 13, 1891, in Villach, Austria, and died in Ann Arbor, Michigan, on January 2, 1992. Kurath came to the United States in 1907, becoming an American citizen in 1912. He was educated at the German-American Teachers Seminar in Milwaukee, Wisconsin, and the Universities of Wisconsin, Texas (A.B. 1914), and Chicago (Ph.D. 1920). Kurath taught German at the University of Texas from 1914 to 1916, and from 1917 to 1919, and taught from 1920 to 1927 at Northwestern University. In 1927 he became professor of German and linguistics at Ohio State University, and in 1930 he was named director of the Linguistic Atlas of the United States. He left Ohio State for Brown University in 1931, where he served as professor of Germanics and linguistics, as well as chair of the department of German and the division of modern languages. In 1946, he became professor of English and editor of the Middle English Dictionary (hereafter MED) at the University of Michigan. Kurath was president of the Linguistic Society of America in 1942 and directed the society's summer Linguistic Institutes from 1947 to 1950. He was the recipient of various scholarly honors; he received honorary doctorates from the University of Chicago and the University of Wisconsin at Milwaukee, as well as the University of Michigan Press Book Award in 1972 for the MED (which he shared with his successors as editor). The esteem his students and colleagues held him in is evident from the volume of papers dedicated to him (Scholler and Reidy, 1973). Kurath retired from Michigan in 1962 but remained scholarly active until his death.

Kurath had a major effect on two fields of linguistics: dialectology and lexicography. His dialect work was rigorously based on Neogrammarian ideas about sound change and resembled similar European projects in issues like choice of linguistic consultants. In addition to directing the Linguistic Atlas project, Kurath published numerous groundbreaking works in this area (e.g., Kurath [1972] and Kurath and McDavid [1961]), a number of his fieldworkers later became prominent dialectologists in their own rights, and the data collected under his direction continue to be useful to dialectologists and sociolinguists. His lexicographical work is also substantial. Ironically, given the overwhelmingly positive reception of the MED, there was some resistance to his

appointment as director of the project, as he was not trained as a medievalist or as a lexicographer and, in fact, published nothing on medieval subjects except for the MED (Bailey, 1992: 806). However, Kurath was able to overcome this barrier by applying the same rigor to his lexicographic work as he did to his dialectological work, and the results speak for themselves.

See also: English, Middle English; English Lexicography; Middle English Dictionaries.

Bibliography

Bailey R W (1992). Hans Kurath. *Language 68,* 797–808.
Kurath H (1972). *Studies in area linguistics.* Bloomington: Indiana University Press.
Kurath H & McDavid R I Jr. (1961). *The pronunciation of English in the Atlantic states.* Ann Arbor: University of Michigan Press.
McDavid R I Jr (1960). Hans Kurath. *Orbis 9,* 597–610.
Scholler H & Reidy J (eds.) (1973). *Lexicography and dialect geography. Festgabe für Hans Kurath.* Wiesbaden: Franz Steiner.

Kurdish

P O Skjærvø, Harvard University, Cambridge, MA, USA

The Kurdish language belongs to the Iranian language family. It is spoken mainly in eastern Turkey, Syria, northern Iraq, western Iran, and Central Asia. Today there are large communities of Kurds living in the diaspora, for instance, Germany and Scandinavia. Kurdish is spoken in three main variants: Northern Kurdish, comprising Kurmanji in the west and dialects spoken from Armenia to Kazakhstan; Central Kurdish, spoken in northeastern Iraq (called Sorani) and adjacent areas in Iran (called Kordi or Mokri), as well as in Iranian Kurdistan (called Senne'i); and Southern Kurdish, spoken in Kermanshah province in western Iran (including Lakki and Lori of Posht-e Kuh). Northern and Central Kurdish developed rich literatures from the early 20th century on.

The earliest grammar and vocabulary of Kurdish were prepared by the Catholic missionary Maurizio Garzoni and published in 1787; these were used in subsequent scholarly descriptions of Kurdish. The earliest modern 20th-century studies were those of Oskar Mann and Karl Hadank. The first important post-World War II study, applying modern linguistic methods, was that of D. Neil MacKenzie (1961–1962).

Literary Kurdish in Iraq, Iran, and the former Soviet Union is written in the Arabo-Persian script, but employs a circumflex accent ^ placed above or below letters to mark non-Persian sounds: above *w* and *y* it denotes *majhul* vowels (*ō*, *ē*, contrasting with *ū* and *ī*), above *l* it denotes *ɫ*, and below *r* it denotes the rolled *r* as opposed to the single-flap *r*. Kurdish in the USSR was also written in the Cyrillic script, with the addition of several signs and diacritics. Kurmanji Kurdish is today written in the Latin alphabet with Turkish orthography, thus <c> = ǰ, <ç> = *č*, <ş> = *š*, <j> = *ž*. In addition, a circumflex accent denotes long vowels and an umlaut on <x> denotes the voiced ɣ. In this script, <e> and <a> represent the vowels commonly transcribed as *a* [a > ɛ] and *ā* [aː > *a*ː].

Kurdish belongs to the Central Iranian language group and, as such, has *s*, *z* from **ć*, *ź* (e.g., *āsin* 'iron', but Persian [Farsi] *āhan*; NK *az* 'I', but Old Persian *adam*). Like the Central dialects and, for instance, Parthian, it has -*ž*- from *-*č*-, -*ǰ*- (*rož* ← **raučah* 'day', Persian *ruz*; *dirēž* ← **drāǰah* 'long', Persian *derāz*). It has also retained the Middle Iranian *majhul* vowels *ō*, *ē*, which in Persian have merged with *ū* and *ī*. A rare feature is the development of intervocalic *m* > *v* (including *m* < *hm* < *šm*), which in Northern and Central Kurdish remains distinct from *w*, but in Southern Kurdish merges with it (demonstrative pronouns: NK *av* 'this' ~ *aw* 'that' ← **ima*- ~ **awa*-; CK both *aw* except an area with *am* ~ *aw*; SK *ī* (cf. Persian *īn*) ~ *ow*; NK *čāv*, CK *čāw* 'eye'; CK *awa* ← **ašmā* 'you'). Kurdish shares with Persian the development of **w*- → *b*- (*bā* 'wind' ← **wātah*, Persian *bād*). Northern Kurdish retains final -*t*, which elsewhere becomes *d* or is lost (*dīt* 'saw', Persian *did*).

The Kurdish dialects have very complex phonologies, morphologies, morphophonologies, and syntax, of which no idea can be given in a small space. Proto-Kurdish had two genders, masculine and feminine, two numbers, and two cases, direct and oblique, as well as a vocative. These are preserved in Northern Kurdish, but gradually merge into a no-gender, no-case system as one moves southward (cf. the 1st sing. personal pronoun 'I, me': NK DIR *az* ~ OBL *min*, CK *(a)min*, SK *amin*). All three groups have an indefinite suffix going back to *-*ēk* 'one', while Central and Southern Kurdish have a definite suffix going back

appointment as director of the project, as he was not trained as a medievalist or as a lexicographer and, in fact, published nothing on medieval subjects except for the MED (Bailey, 1992: 806). However, Kurath was able to overcome this barrier by applying the same rigor to his lexicographic work as he did to his dialectological work, and the results speak for themselves.

See also: English, Middle English; English Lexicography; Middle English Dictionaries.

Bibliography

Bailey R W (1992). Hans Kurath. *Language 68,* 797–808.

Kurath H (1972). *Studies in area linguistics.* Bloomington: Indiana University Press.

Kurath H & McDavid R I Jr. (1961). *The pronunciation of English in the Atlantic states.* Ann Arbor: University of Michigan Press.

McDavid R I Jr (1960). Hans Kurath. *Orbis 9,* 597–610.

Scholler H & Reidy J (eds.) (1973). *Lexicography and dialect geography. Festgabe für Hans Kurath.* Wiesbaden: Franz Steiner.

Kurdish

P O Skjærvø, Harvard University, Cambridge, MA, USA

The Kurdish language belongs to the Iranian language family. It is spoken mainly in eastern Turkey, Syria, northern Iraq, western Iran, and Central Asia. Today there are large communities of Kurds living in the diaspora, for instance, Germany and Scandinavia. Kurdish is spoken in three main variants: Northern Kurdish, comprising Kurmanji in the west and dialects spoken from Armenia to Kazakhstan; Central Kurdish, spoken in northeastern Iraq (called Sorani) and adjacent areas in Iran (called Kordi or Mokri), as well as in Iranian Kurdistan (called Senne'i); and Southern Kurdish, spoken in Kermanshah province in western Iran (including Lakki and Lori of Posht-e Kuh). Northern and Central Kurdish developed rich literatures from the early 20th century on.

The earliest grammar and vocabulary of Kurdish were prepared by the Catholic missionary Maurizio Garzoni and published in 1787; these were used in subsequent scholarly descriptions of Kurdish. The earliest modern 20th-century studies were those of Oskar Mann and Karl Hadank. The first important post-World War II study, applying modern linguistic methods, was that of D. Neil MacKenzie (1961–1962).

Literary Kurdish in Iraq, Iran, and the former Soviet Union is written in the Arabo-Persian script, but employs a circumflex accent ^ placed above or below letters to mark non-Persian sounds: above *w* and *y* it denotes *majhul* vowels (*ō*, *ē*, contrasting with *ū* and *ī*), above *l* it denotes *ł*, and below *r* it denotes the rolled *r* as opposed to the single-flap *r*. Kurdish in the USSR was also written in the Cyrillic script, with the addition of several signs and diacritics. Kurmanji Kurdish is today written in the Latin alphabet with Turkish orthography, thus <c> = ǰ, <ç> = č, <ş> = š, <j> = ž. In addition, a circumflex accent denotes long vowels and an umlaut on <x> denotes the voiced ɣ. In this script, <e> and <a> represent the vowels commonly transcribed as *a* [a > ɛ] and *ā* [a: > *a*:].

Kurdish belongs to the Central Iranian language group and, as such, has *s*, *z* from **ć*, *ȷ́* (e.g., *āsin* 'iron', but Persian [Farsi] *āhan*; NK *az* 'I', but Old Persian *adam*). Like the Central dialects and, for instance, Parthian, it has *-ž-* from **-č-*, *-ǰ-* (*rož* ← **raučah* 'day', Persian *ruz*; *dirēž* ← **drāǰah* 'long', Persian *derāz*). It has also retained the Middle Iranian *majhul* vowels *ō*, *ē*, which in Persian have merged with *ū* and *ī*. A rare feature is the development of intervocalic *m* > *v* (including *m* < *hm* < *šm*), which in Northern and Central Kurdish remains distinct from *w*, but in Southern Kurdish merges with it (demonstrative pronouns: NK *av* 'this' ~ *aw* 'that' ← **ima-* ~ **awa-*; CK both *aw* except an area with *am* ~ *aw*; SK *ī* (cf. Persian *īn*) ~ *ow*; NK *čāv*, CK *čāw* 'eye'; CK *awa* ← **ašmā* 'you'). Kurdish shares with Persian the development of **w-* → *b-* (*bā* 'wind' ← **wātah*, Persian *bād*). Northern Kurdish retains final *-t*, which elsewhere becomes *d* or is lost (*dīt* 'saw', Persian *did*).

The Kurdish dialects have very complex phonologies, morphologies, morphophonologies, and syntax, of which no idea can be given in a small space. Proto-Kurdish had two genders, masculine and feminine, two numbers, and two cases, direct and oblique, as well as a vocative. These are preserved in Northern Kurdish, but gradually merge into a no-gender, no-case system as one moves southward (cf. the 1st sing. personal pronoun 'I, me': NK DIR *az* ~ OBL *min*, CK *(a)min*, SK *amin*). All three groups have an indefinite suffix going back to **-ēk* 'one', while Central and Southern Kurdish have a definite suffix going back

to *-*aka* (-*ak*). The 3rd singular enclitic pronoun is -*ī*, pl. -*yān* (cf. Persian -*eš*, -*ešān*). The 1st and 2nd plural enclitic pronouns are archaic: 1PL -*n*, 2PL -*ū* (beside -*mān* and -*tān*), which go back to Old Iranian **nah* and **wah*. The ezafe has two genders and two numbers, but is simplified according to the general tendency.

The verbal systems are of the typical modern Iranian type, with a three-stem system: present-imperfect, past, and perfect, as well as a split ergative. The present and imperfect take the prefixes NK *t-/di-*, CK *a-/da-*, SK *a-/mi-* (cf. Persian *mi-*) to express progressive tense. Present forms without a prefix or with the prefix *bi-* are subjunctive. The past tenses are ergativic, but only Northern Kurdish has the pure passive type construction (1).

(1) ta az nās kir-im
you.SING.OBL I.DIR familiar do.PAST-1ST.SING
'you knew me'

In Central and Southern Kurdish, where there is no case distinction, the agent is expressed by enclitic pronouns. If the verb has a preverb, the enclitic pronouns come after it (2), otherwise they come between the verb stem and the ending (3).

(2) a-t xwārd
PROG-you.AGT eat.PAST[-3SG]
'you were eating'

(3) nārd-it-īn
send.PAST-you.AGENT-COP.1PL
'you sent us'

Note constructions of the type in (4), where the apposition *pēwa* 'about' governs the person implicit in the copula -*īt* 'you are'; (5), where 'to' governs the pronoun implicit in the copula -*n*- 'you are'; and (6), where 'to' governs the pronoun implicit in the copula -*m*- 'I am'.

(4) xaw-im pēwa dīw-īt
dream-1SG.ENCL about see.PAST-COP.2SG
'I dreamed about you'

(5) dā-m-ī-n-ē
give.PAST-COP.1SG-he.ENCL.OBL-cop.2PL-to
'he gave me to you'

(6) kitēb-ek-ān-it dā-m-ē
book-DEF-PLUR-you.AGT give.PAST-COP.1SG-to
'you gave me the books'

In Northern Kurdish, the passive is formed with the auxiliary *hāt*-'come' + INF (7).

(7) hāt-iye girt-in
come.PERF[-3SG] seize.PAST.INF
'has been seized'

Central and Southern Kurdish have passive formations in present *rē-* or *yē-*, past-*rā-* or -*yā-* (*kuž-yā-* 'be killed'; *de-nūs-rē* [PROG-write.PRES-PASS.PRES. [-3SG]] 'it is being written').

Derivational nominal suffixes are common, as are a variety of types of compound. The meaning of verbs can be modified by preverbs, of which there are many, or the postverb -*ava-/-awa* (-*āw*), unique to Kurdish among Iranian dialects (*hāt-in* [come.INF] 'to come', but *hāt-in-āw* 'to come back, return'. Verbal idioms consisting of adjectives or nouns plus verbs are common, as in all Iranian languages.

See also: Armenia: Language Situation; Iranian Languages; Iran: Language Situation; Kazakhstan: Language Situation; Persian, Modern; Turkey: Language Situation.

Bibliography

Blau J (1989). 'Le kurde.' In Schmitt R (ed.) *Compendium linguarum iranicarum.* Wiesbaden: Reichert. 327–335.

Blau J (2000). *Méthode de kurde: sorani* (rev. edn.). Paris: Harmattan.

Blau J & Barak V (1999). *Manuel de kurde: kurmanji.* Paris: Harmattan.

MacKenzie D N (1961). 'The origins of Kurdish.' *Transactions of the Philological Society 59,* 68–86.

MacKenzie D N (1961–1962). *Kurdish dialect studies* (2 vols.). London/New York: Oxford University Press.

MacKenzie D N (1986). 'Kurds, Kurdestan v. language.' In Bearman P J, Bianquis T, Bosworth C E, van Donzel E & Heinrichs W P (eds.) *Encyclopædia of Islam,* new edn., vol. 5. Leiden: Brill. 479–480.

Kurukh

P S Subrahmanyam, Annamalai University, Bangalore, India

Introduction

Kuṛux is one of the tribal languages of the Dravidian family with a population of about 1.5 million; thus it is next only to Gondi within the family in terms of number of speakers. The name of the language is spelt also as Kurukh; another name, Oraon, was also in use in earlier times. The main concentration of Kuṛux speakers is found in Chota Nagpur and Bhagalpur districts of Bihar, India; some of them live also in Madhya Pradesh (Raygarh and Sarguja districts) and Orissa (Sundargarh and Sambalpur districts). Some have migrated in recent times to the tea districts of Assam and Nepal and are known there as Dhāngar/Dhangar 'men who receive dhān "rice" as wages'. It belongs to the North Dravidian subgroup along with Malto, its closest ally, and Brahui.

Phonology

Kuṛux contains the ten-vowel system that is normally found in the Dravidian languages (see Table 1); there are also a few words with nasalized vowels, e.g., *khẽ:sō* 'blood'. Its consonant system is presented in Table 2. The velar voiceless fricative (x) and the glottal stop (ʔ) are the peculiar features of Kuṛux phonology, e.g., *xay* 'wife', *ciʔinā* 'to give'.

Syntax

Word Classes

The following word classes may be recognized for Kuṛux: nouns (including pronouns and numerals), verbs, adjectives, adverbs (including expressives), particles, and interjections.

The class of adjectives is a small one. An adjective occurs before the noun it qualifies. Most of the nouns can function as adjectives, for example:

mechā partā
high, height *mountain*
'high mountain'

The adjectives derived from the three deictic bases (see 'Pronouns') show agreement for number (but not for gender) with the noun that follows. Thus, *ā* 'that', *ī* 'this' and *ū* 'that at a greater distance' are used with a singular noun, e.g., *ā/ī/ū ālas* 'that/this/that (extra-dist.) man'. The corresponding forms used with a plural noun are *abṛā*, *ibṛā*, and *hubṛā*, e.g., *abṛā/ibṛā/hubṛā āl-ar* 'those/these/those (extra-dist.) men'. (For verbal adjectives, see 'Nonfinite Verbs').

An adverb occurs before the verb, for example:

ād xanem xanem bar-ckī raʔī
she again again come-past was
'She came frequently.'

Adverbs may be divided into those of (a) time (e.g., *cērō* 'yesterday', *innā* 'today', *nēlā* 'tomorrow'), (b) place (e.g., *mund* 'before, in front', *iyā* 'here', *ayā* 'there', *eksan* 'where'), and (c) manner (e.g., *baggī/baggū* 'much', *oṇṭā oṇṭā* 'separately, one by one', *dav/davdim* 'well').

The particle of emphasis is a good example for particles. It is +*m/im/am/em*, e.g., *nēla+m* 'tomorrow itself' (: *nēlā* 'tomorrow'), *ēn-im* 'even I', *ār-im* 'even they'.

Examples for interjections include: *haʔī* 'yes', *mal/malʔā/mallā* 'no', *anti(jē)* 'of course'.

Word Order

The favored word order in Kuṛux is S(ubject) O(bject) V(erb):

ās eŋg-āgē mandar ci-c-as
he I-DAT medicine give-PAST-3M.SG
'He gave me medicine.'

Table 1 Vowels of Kuṛux

	Front		*Central*		*Back*	
	Short	*Long*	*Short*	*Long*	*Short*	*Long*
High	i	ī			u	ū
Mid	e	ē			o	ōō
Low			a	ā		

Table 2 Consonants of Kuṛux

	L	*D*	*R*	*P*	*Vel*	*G*
Stop						
VL	p	t	ṭ	c	k	ʔ
VLA	ph	th	ṭh	ch	kh	
VD	b	d	ḍ	j	g	
VA	bh	dh	ḍh	jh	gh	
Nasal	m	n			ŋ	
Fricative		s			x	h
Lateral		l				
Trill		r				
Flap			ṛ			
Semivowel	w			y		

(Abbreviations: D = dental, G = glottal, L = labial, P = palatal, R = retroflex, VD = voiced (unaspirated), VA = voiced aspirated, Vel = velar, VL = voiceless (unaspirated), VLA = voiceless aspirated.)

Gender and Number

Kuṛux shows a two-way distinction in gender but the classification differs in the singular and the plural; the feminine goes with the nonhuman in the singular but with the masculine in the plural, as illustrated by the following pronouns:

ās
'he'
ād
'she, it'
ār
'they (human)'
abṛā
'they (nonhuman)'.

Agreement

The finite verb shows agreement with the subject pronoun (or a corresponding noun in the case of the third person) by a change in the personal suffix (the final morpheme in the forms shown in Table 3). In equational sentences, the predicate noun shows agreement with the subject pronoun by taking the same personal suffixes as the finite verb (except 3rd non-m. sg./non-hum. pl. [see 'Finite Verbs']). The use of the copular verb (*tal-* 'to be') in such sentences is optional. Kuṛux partially preserves the old Dravidian feature of the absence of the copular verb:

ēn kuṛux-an (tal-d-an)
I Kuṛux-1sg (be-pres-1sg)
'I am a Kuṛux (speaker).'
ēm kuṛux-am (tal-d-am)
we (excl.) Kuṛux-1pl.excl (be-pres-1pl.excl)
'We (excl.) are Kuṛux (speakers).'

For agreement between the demonstrative adjectives and the nouns qualified with regard to number, see 'Word Classes.'

Table 3 The finite forms of ijʔ- 'to stand'

	Pronoun	*Past tense*	*Present tense*	*Future tense*
1sg.	ēn	ij-k-an	ij-d-an	ijʔ-o-n
1pl. (excl.)	ēm	ij-k-am	ij-d-am	ijʔ-o-m
1pl. (incl.)	nām	ij-k-at	ij-d-at	ijʔ-o-t
2sg.	nīn	ij-k-ay	ij-d-ay	ijʔ-o-y
2sg. (f.)		ij-k-i	ij-d-i	ijʔ-o-y
2pl.	nīm	ij-k-ar	ij-d-ar	ijʔ-o-r
3m.sg	ās	ijj-Ø-as	ij-d-as	ijʔ-o-s
3h.pl.	ār	ijj-Ø-ar	ij-n-ar	ijʔ-o-r
3non-m.sg./ non-h.pl.	abṛā	ijj-Ø-a	ijʔ-i	ijʔ-o-Ø

Noun Morphology

A nominal base is followed by the plural suffix when plurality has to be expressed; the case suffix/postposition occurs at the end. A masculine noun may take the definite suffix *-as*, e.g., *āl* 'man', *āl-as* 'a particular man'.

Plural Suffixes

The plural suffix for the human nouns is *-ar*, e.g., *āl-ar* 'men', *mukk-ar* 'women' (sg. *mukkā*). Some feminine nouns take *-guṭhi-ar*, e.g., *ālī-guṭhi-ar* 'wives'; *-guṭhi (-ar)* is also added optionally to human nouns with *-ar* plural, e.g., *āl-ar(-guṭhi(-ar))* 'men'. Kinship terms, however, take *-baggar*, e.g., *dādā-baggar* 'elder brothers'. The word *xadd-xarrā* 'children' seems to contain another plural suffix *-xarrā*. Nonhuman nouns optionally add *-guṭhī*, which in origin is a separate word meaning 'group', to form the plural, for example:

mūnd aḍḍō(-guṭhī)
three OX-PL
'three oxen'
enḍ man(-guṭhī)
two *tree*-PL
'two trees'.

Case Suffixes and Postpositions

The nominative is unmarked. When a noun is used in the vocative, *-ō/-ay(ō)* is added to it at the end; the vocative form may be preceded by *ē/ana* (but *anay* when a woman is addressed), e.g., *ē/ana urb-ay(ō)* 'O master!', *anay mukk-ay* 'O woman!'. Further, when women talk to women, *anay* is replaced by *ān/anē*, e.g., *ān xay* 'O daughter!', *anē xay-guṭhi-ar-ō* 'O daughters!'.

The accusative suffix is *-an* after a consonant and *-n* after a vowel, e.g., *āl-an* 'man (accus.)', *mukka-n* 'woman (accus.)' (: *mukkā* 'woman') but it is *-in* after the definite suffix *-as* (in masculine nouns, e.g., *āl-as-in* 'the man (accus.)', the plural *-ar* (e.g., *āl-ar-in* 'men (accus.)') and after a demonstrative pronoun that ends in a consonant, e.g., *ād-in* 'her, it (accus.)'.

The instrumental suffix is *-trī/-trū*, e.g., *āl-trī* 'by the man'.

The dative uses the postposition *+gē* (variant *+ā(gē)* after a pronoun), e.g., *āl+gē* 'to the man', *eŋg-ā(gē)* 'to me'.

The ablative suffix is *-tī* after a consonant, *-ntī* after a vowel; after the plural *-ar*, *-tī* freely varies with *-intī*, e.g., *āl-tī* 'from the man', *mukka-ntī* 'from the woman', *āl-ar-tī* /*āl-ar-intī* 'from the men'.

The genitive commonly uses the postposition *+gahi*, e.g., *āl + gahi* 'of the man'; it has the variant

+*hay* after a pronoun, e.g., *eŋg* + *hay* 'my'. However, nouns denoting a place take the suffix *-ntā*, e.g., *padda-ntā* 'of the village' (: *paddā* 'village').
The locative uses the postposition +*nū* (also +*nō* dialectally), e.g., *eṛpā* + *nū* 'in the house'.

Pronouns

Two important features of the pronominal system of Kuṛux are the presence of inclusive and exclusive distinction in the first person pronoun (represented also in the finite verb) and the formation of the third person pronouns on three deictic bases.
The personal pronouns are:

ēn 'I'	ēm 'we (excl.)'
	nām 'we (incl.)'
nīn 'you (sg.)'	nīm 'you (pl.)'

The third person pronouns are formed on the three deictic bases *a-/ā-* (distant), *i-/ī-* (proximate) and *hu-/hū-* (extra-distant):

Distant	Proximate	Extra-distant	
ā-s	ī-s	hū-s	'he'
ā-d	ī-d	hū-d	'she, it'
ā-r	ī-r	hū-r	'they (hum.)'
a-bṛā	i-bṛā	hu-bṛā	'they (non-hum.)'

The reflexive pronouns are *tān* (sg.) and *tām* (pl.).
There are five interrogative pronouns: *nē* 'who', *ēkā* 'who, which', *endr/endrā/ēkdā* 'what, which'. The addition of the particle +*ʔim/ʔam* to an interrogative pronoun converts it into an indefinite pronoun, e.g., *nīk+ʔim* 'someone'.

Numerals

Kuṛux retains the Dravidian numerals only up to 4 and the rest are borrowed from Hindi. The native numerals have also preserved a distinction between nonhuman and human (the variants given in parentheses for the human forms occur before case suffixes and postpositions).

	Nonhuman	Human
'1'	oṇḍ/oṇṭā	ort
'2'	eṇḍ/e:ṇḍ	irb (irbar-)
'3'	mūnd	nub (nubar-)
'4'	nāx	naib (naibar-)

The human numerals are generally followed by the classifier *jhan-ar* 'persons' (sg. *jhan*), for example:

nub	jhan-ar	bar-c-ar
three (h.)	CLASSIF	*come*-PAST-H.PL.

'Three men came.'

The counting is done in terms of 'score' for which the word is *kūrī/bīsoʔe*, *oṇḍ kūrī/bīsoʔe* 'one score', *kūrīeṇḍ* or *eṇḍ bīsoʔe* 'two score', etc. The ordinal is formed by adding *-(an)-tā* to a cardinal, e.g., *eṇḍ-(an)tā* 'second'; however, there is a special word for 'first': *mund-(an)tā*.

Verb Morphology

Female Speech

An important characteristic feature of the Kuṛux verb morphology is the use of special forms by women when they talk among themselves; some of the tense and personal suffixes in women's speech are different from those in men's speech:

(Men's speech)	nīn	ekātarā	kaʔa-d-ay
(Women's speech)	nīn	ekātarā	kaʔa-d-ī
	you (sg.)	*where*	*go*-PRES-2SG

'Where are you (sg.) going?'

Verb Bases

A verb base in Kuṛux may be simple or complex. Four types may be recognized among the complex bases:

1. Intransitives or reflexives derived from transitives by the addition of *-r-*, e.g., *kol-r-* 'to be opened' (: *kol-* 'to open'), *mūjh-r-* 'to wash one's own face' (: *mūjh-* 'to wash another's face').
2. Transitives derived from intransitives by the addition of *-aʔa-*, e.g., *maṛx-aʔa-* 'to make dirty' (: *maṛx-* 'to be dirty').
3. Causatives derived from transitives by the addition of *-tāʔa-*, e.g., *es-tāʔa-* 'to cause to break' (: *ess-*'to break').
4. Reciprocals derived by the addition of *-nakr-* to a simple verb base, e.g., *ēr-nakr-* 'to look at each other' (: *ēr-* 'to see'), *kēb-nakr-* 'to abuse one another' (: *kēb-* 'to abuse').

Finite Verbs

A finite verb has the following structure:

Verb Base + Tense Suffix + Personal Suffix

The personal suffixes are:

1st sg.	-an
1st pl. (excl.)	-am
1st pl. (incl.)	-at
2nd sg.	-ay /(female speech) -i
2nd pl.	-ar
3rd m. sg.	-as
3rd hum. pl.	-ar
3rd non-m. sg./ non-hum. pl.	-a (after past)/ -i (after present)

The vowel of all these suffixes is deleted after the future suffix *-o*. Three tenses, past (suffixes: *-k-*,*-ck-*), present (suffixes: *-d-*, *-n-*) and future (*-o-*) are there (see Table 3).

Unlike its counterpart in the sister languages, the imperative in Kuṛux has only one form without distinction between the singular and the plural. The suffix is *-ā* when men are addressed and *-ay* when women are addressed (when women speak to women, the suffix is *-ē*): *esʔ-ā/-ay/-ē* 'Break!'. A milder sort of the imperative has the suffix *-kē*, e.g., *bar-kē* 'Come (if you please)!'.

Nonfinite Verbs

The infinitive, which can also function as a noun and an adjective, has the suffix *-nā* (also *-ā* in certain constructions), e.g., *es-nā/esʔ-ā* 'to break, breaking'.

The present participle has the suffix *-nū(tī)/-num*, e.g., *es-nū(tī)/es-num* 'breaking'.

The past participle has the suffix *-ār*, e.g., *esʔ-ār* 'having broken'.

There are three types of adjectives derived from verb bases:

- The infinitive with adjectival function:

kūr-nā	amm
burning	*water*

'hot water'

- Past adjective with the suffix *-ckā*:

ke-ckā	āl-ar
die-PAST ADJ	*man-PL*

'dead people'

- Agent noun (see *Agent Noun*) functioning as nonpast adjective:

pāṛ-ū/ pāṛ-nā	pellō
sing-NON-PAST.ADJ	*girl*

'singing girl'

Agent Noun

The agent noun, with the suffix *-ū* can also take the masculine and the plural suffixes, e.g., (from *esʔ-* 'to break')

isʔ-ū	'one who/which breaks'
isʔ-u-s	'a man who breaks'
isʔ-u-r	'persons who break'

Negative Verb

Kuṛux employs the verb base *mal-* 'to be not' in the present tense (with *-y-* and *-k-* as optional variants of the regular present suffixes *-d-/-n-* [see Table 3]) as the negative verb to deny the identity of the subject and the predicate noun phrases, for example:

ēn	bēl-an	mal-d-an/mal-y-an/mal-k-an
I	*king-1SG*	*not.be-PRES-1SG*

'I am not a king.'

See also: Dravidian Languages; India: Language Situation.

Bibliography

Bhat D N S (1966). 'Kuṛux nouns.' *Indian Linguistics 27,* 102–108.

Bhat D N S (1970a). 'Kuṛux indicatives.' *Indo-Iranian Journal 12,* 216–223.

Bhat D N S (1970b). 'Velar consonants in Kuṛux.' *Indian Linguistics 31,* 177–181.

Grignard A (1924a). *A grammar of the Oraon language.* Calcutta: Catholic Orphan Press.

Grignard A (1924b). *An Oraon-English dictionary.* Calcutta and Vienna.

Hahn F (1900). *Kurukh grammar.* Calcutta: Bengal Secretariat Press. (Repr. Delhī Mittal Publications [no date].)

Hahn F (1903). *Kurukh (Orāō"*)*-English dictionary.* Part I. Calcutta.

Pfeiffer M (1972). *Elements of Kuṛux historical phonology.* Leiden: E. J. Brill.

Kuryłowicz, Jerzy (1895–1978)

R D Greenberg, Yale University, New Haven, CT, USA

Jerzy Kuryłowicz was born in the city of Stanislau in 1895, which at that time was part of the Austro–Hungarian Empire, and is now known as Ivano–Frankivs'k in Ukraine. He studied foreign languages while he was both a secondary school student and while studying for his first university degree in Vienna. His university studies were interrupted by the outbreak of World War I, when Kuryłowicz was drafted for military service. He was injured in the war and spent some time recovering from his injuries in a Russian hospital, where he is said to have perfected his knowledge of Russian; he also became interested in language loss among aphasia patients. In 1920, he pursued graduate studies in Romance and Germanic

Unlike its counterpart in the sister languages, the imperative in Kuṛux has only one form without distinction between the singular and the plural. The suffix is *-ā* when men are addressed and *-ay* when women are addressed (when women speak to women, the suffix is *-ē*): *esʔ-ā/-ay/-ē* ‘Break!’. A milder sort of the imperative has the suffix *-kē*, e.g., *bar-kē* ‘Come (if you please)!’.

Nonfinite Verbs

The infinitive, which can also function as a noun and an adjective, has the suffix *-nā* (also *-ā* in certain constructions), e.g., *es-nā/esʔ-ā* ‘to break, breaking’.

The present participle has the suffix *-nū(tī)/-num*, e.g., *es-nū(tī)/es-num* ‘breaking’.

The past participle has the suffix *-ār*, e.g., *esʔ-ār* ‘having broken’.

There are three types of adjectives derived from verb bases:

- The infinitive with adjectival function:

kūr-nā	amm
burning	*water*

‘hot water’

- Past adjective with the suffix *-ckā*:

ke-ckā	āl-ar
die-PAST ADJ	*man*-PL

‘dead people’

- Agent noun (see *Agent Noun*) functioning as nonpast adjective:

pāṛ-ū/ pāṛ-nā	pellō
sing-NON-PAST.ADJ	*girl*

‘singing girl’

Agent Noun

The agent noun, with the suffix *-ū* can also take the masculine and the plural suffixes, e.g., (from *esʔ-* ‘to break’)

isʔ-ū	‘one who/which breaks’
isʔ-u-s	‘a man who breaks’
isʔ-u-r	‘persons who break’

Negative Verb

Kuṛux employs the verb base *mal-* ‘to be not’ in the present tense (with *-y-* and *-k-* as optional variants of the regular present suffixes *-d-/-n-* [see **Table 3**]) as the negative verb to deny the identity of the subject and the predicate noun phrases, for example:

ēn	bēl-an	mal-d-an/mal-y-an/mal-k-an
I	*king*-1SG	*not.be*-PRES-1SG

‘I am not a king.’

See also: Dravidian Languages; India: Language Situation.

Bibliography

Bhat D N S (1966). ‘Kuṛux nouns.’ *Indian Linguistics 27,* 102–108.
Bhat D N S (1970a). ‘Kuṛux indicatives.’ *Indo-Iranian Journal 12,* 216–223.
Bhat D N S (1970b). ‘Velar consonants in Kuṛux.’ *Indian Linguistics 31,* 177–181.
Grignard A (1924a). *A grammar of the Oraon language.* Calcutta: Catholic Orphan Press.
Grignard A (1924b). *An Oraon-English dictionary.* Calcutta and Vienna.
Hahn F (1900). *Kurukh grammar.* Calcutta: Bengal Secretariat Press. (Repr. Delhī Mittal Publications [no date].)
Hahn F (1903). *Kurukh (Orāōⁿ)-English dictionary.* Part I. Calcutta.
Pfeiffer M (1972). *Elements of Kuṛux historical phonology.* Leiden: E. J. Brill.

Kuryłowicz, Jerzy (1895–1978)

R D Greenberg, Yale University, New Haven, CT, USA

Jerzy Kuryłowicz was born in the city of Stanislau in 1895, which at that time was part of the Austro–Hungarian Empire, and is now known as Ivano–Frankivs’k in Ukraine. He studied foreign languages while he was both a secondary school student and while studying for his first university degree in Vienna. His university studies were interrupted by the outbreak of World War I, when Kuryłowicz was drafted for military service. He was injured in the war and spent some time recovering from his injuries in a Russian hospital, where he is said to have perfected his knowledge of Russian; he also became interested in language loss among aphasia patients. In 1920, he pursued graduate studies in Romance and Germanic

philology at the Jan Kazimierz University in Lvov. At that time he also became interested in Indic studies under the tutelage of Andrzej Gawronski, and in Semitic studies with Zygmunt Smogorzewski and Mojzesz Schorr.

He completed his doctorate within three years, and thereafter traveled to Paris on a scholarship from the French government to continue his studies in Romance linguistics. In Paris he attended lectures of the eminent Indo–Europeanist, Antoine Meillet, and under Meillet's influence, Kuryłowicz shifted his main interests from Romance linguistics to comparative Indo–European. In this period, he analyzed the Cuneiform Hittite texts, and through these texts was able to confirm Saussure's 1879 postulations regarding Indo–European laryngeals. This discovery sparked further interest in the study of Indo–European phonology. For his postdoctoral studies in France, Kuryłowicz received a diploma from the *Ecole Pratique des Hautes Etudes*.

Kuryłowicz returned to Poland in 1928 to take up a professorial post at the Jan Kazimierz University of Lvov. In the 1930s, he regularly presented at scholarly linguistic conferences in Europe, and in 1931 received a Rockefeller scholarship, which took him for an academic year to Yale University in the United States. That same year, he was elected a corresponding member of the Polish Academy of Arts and Sciences, and he became a full member of the Academy in 1938. He spent World War II in Lvov during both the Soviet and Nazi occupations. After the war, Lvov became part of the USSR, and Kuryłowicz moved to the devastated city of Wrocław (Breslau before the war) in 1946, where he became a professor at the University of Wrocław. In 1948, he was appointed a professor at the Jagiellonian University in Cracow, where a chair of general linguistics was created for him. He remained in this post until his retirement in 1965.

In his first eight years in Cracow, the entire Polish educational system came under severe pressure due to Stalinization. Departments of English and other west European languages were abolished, and scholars were prohibited from maintaining contacts with or traveling to the West. Nevertheless, this period proved to be a highly productive one for Kuryłowicz. After travel restrictions were relaxed in 1956, Kuryłowicz worked to reestablish departments of English and other Western languages, and rebuilt strong links with linguists in western Europe and the United States. He received several honorary doctorates in the 1960s and 1970s and was a visiting professor at both Harvard University and the Massachusetts Institute of Technology. In 1964, he organized a symposium at the Jagiellonian University in Cracow in honor of the institution's 600th anniversary. The symposium was dedicated to the 'Notion of law in linguistics'; his own paper focused on the laws of isomorphisms. Kuryłowicz remained an active scholar until his death in 1978; he wrote 4 books in retirement and approximately 80 scholarly articles.

Kuryłowicz was a prolific scholar, and his works covered topics in Indo–European, Germanic, Romance, Semitic, and Slavic linguistics, as well as general linguistics. In his early works, he accepted principles of structural linguistics in his analysis of diachronic linguistic material, and he was an admirer of the work of the Prague linguistic circle. In 1948, he brought such a structurally oriented theory of language to Cracow, inspiring a generation of students there, and making Cracow the strongest center for linguistics in Poland. In the interwar period, Kuryłowicz made significant contributions to the study of Indo-European through his work on the Hittite texts. He published his first monograph in 1935, *Etudes indoeuropéennes* ('Indo–European studies'). After World War II, he focused on the accentuation of Indo–European languages and the issue of apophony (ablaut) in Indo–European languages. He became particularly interested in the subject of linguistic analogy in the historical development of languages, and elaborated upon what he considered to be a series of laws for linguistic analogy. In the 1960s, he gained inspiration through his contacts with Roman Jakobson and Morris Halle, who sparked his interest in the subject of inflectional categories. After his retirement, he strove to produce a definitive Indo–European grammar in cooperation with some other linguists. He produced one volume in this series, focusing on Indo–European accentuation and ablaut (1968).

See also: Halle, Morris (b. 1923); Jakobson, Roman (1896–1982); Meillit, Antoine (Paul Jules) (1866–1936); Saussure, Ferdinand (-Mongin) de (1857–1913); Structuralism.

Bibliography

Kuryłowicz J (1935). *Etudes indoeuropéennes ('Indo–European studies')* (vol. 1). Cracow: Gebethner and Wolff.

Kuryłowicz J (1952/1958). *L'Accentuation des langues indoeuropéennes ('The accentuation of the Indo–European languages')*. Cracow: Nakladem PAU (2nd edn. Wrocław and Cracow: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN).

Kuryłowicz J (1956). *L'Apophonie en indo–européen* ('Indo–European apophony'). Wroclaw: Zaklad im. Ossolinskich; Warsaw: Wydawnictwo PAN.

Kuryłowicz J (1960/1973). *Esquisses linguistiques.* ('Linguistic sketches'). Wrocław–Warsaw–Cracow: (2nd edn. Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN Munich: Wilhelm Fink).
Kuryłowicz J (1961). *L'Apophonie en semitique* ('Apophony in Semitic'). Wroclaw–Warsaw–Cracow: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.
Kuryłowicz J (1964). *The inflectional categories of Indo–European.* Heidelberg: Carl Winter.
Kuryłowicz J (1968). *Indogermanische Grammatik* ('Indo–European grammar') (vol. 2). Heidelberg: Carl Winter.
Kuryłowicz J (1972). *Studies in Semitic grammar and metrics.* Wroclaw–Warsaw–Cracow: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.
Kuryłowicz J (1975). *Metrik und Sprachgeschichte* ('Metrics and the history of language'). Wroclaw–Warsaw–Cracow–Gdansk: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.
Kuryłowicz J (1977). *Problèmes de linguistique indo-européenne.* ('Problems of Indo–European linguistics'). Wroclaw–Warsaw–Cracow–Gdnask: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.

Kuwait: Language Situation

Editorial Team

On the northwestern coast of the Persian Gulf, Kuwait shares borders with Iraq and Saudi Arabia. Kuwait's modern history began in the early 18th century, when several clans from the Al Aniza tribe migrated to the northern shore of the Gulf from the Najd, their famine-stricken homeland in central Arabia. Eventually, the Al Sabah emerged as the dominant clan, and its members were formally established as rulers in 1756. By the late 19th century, however, fears of growing Ottoman influence led Sheikh Mubarak Al Sabah to enter into an agreement with Great Britain, which effectively established Kuwait as an autonomous British protectorate. Under the 1899 agreement, Kuwait maintained control over its internal affairs, and Great Britain assumed responsibility for the country's security and foreign relations. By the 1960s, Kuwait enjoyed such a level of economic development that it became one of the wealthiest states on a per capita basis, and in 1961 Kuwait gained full independence from Britain. For a brief period, between 1990 and 1991, Kuwait was under Iraqi control, but the Iraqi army was defeated by United Nation forces.

The official national language of Kuwait is Standard Arabic, a modernized variety of Classical Arabic. The latter, with its archaic vocabulary, is used for religious and ceremonial purposes, whereas Standard Arabic is employed in formal situations, for education, and for official purposes and is the language of the press. In the Arabic-speaking world, only the well educated have adequate proficiency in Standard Arabic; in Kuwait, the literacy rate among men is 85.1% and among women is 81.7%. About 85% of the population uses Gulf Spoken Arabic to communicate. Mehri, a South Arabian language, is used by some 3530 Kuwaitis. In addition, English is widely spoken, and both television and radio broadcasts are often in English. There is one institution in Kuwait for all its 110 285 deaf inhabitants.

See also: Arabic.

Kwa Languages

M E K Dakubu, University of Ghana, Accra, Ghana

The Kwa languages, a branch of the Niger-Congo family, are spoken in West Africa. As presently defined, the group extends from southeastern Ivory Coast in the west through the southern two-thirds of Ghana, southern Togo, and the Republic of Benin to the Benin-Nigeria border in the east. The term Kwa was adopted by linguists toward the end of the 19th century to group together Akan, Ga, Ewe (Éwé), and their close relatives because in many of these languages a stem *kwa* or *kua* means 'man' or 'person.'

Kuryłowicz J (1960/1973). *Esquisses linguistiques.* ('Linguistic sketches'). Wrocław–Warsaw–Cracow: (2nd edn. Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN Munich: Wilhelm Fink).
Kuryłowicz J (1961). *L'Apophonie en semitique* ('Apophony in Semitic'). Wroclaw–Warsaw–Cracow: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.
Kuryłowicz J (1964). *The inflectional categories of Indo–European.* Heidelberg: Carl Winter.
Kuryłowicz J (1968). *Indogermanische Grammatik* ('Indo–European grammar') (vol. 2). Heidelberg: Carl Winter.
Kuryłowicz J (1972). *Studies in Semitic grammar and metrics.* Wroclaw–Warsaw–Cracow: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.
Kuryłowicz J (1975). *Metrik und Sprachgeschichte* ('Metrics and the history of language'). Wroclaw–Warsaw–Cracow–Gdansk: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.
Kuryłowicz J (1977). *Problèmes de linguistique indo-européenne.* ('Problems of Indo–European linguistics'). Wroclaw–Warsaw–Cracow–Gdnask: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.

Kuwait: Language Situation

Editorial Team

On the northwestern coast of the Persian Gulf, Kuwait shares borders with Iraq and Saudi Arabia. Kuwait's modern history began in the early 18th century, when several clans from the Al Aniza tribe migrated to the northern shore of the Gulf from the Najd, their famine-stricken homeland in central Arabia. Eventually, the Al Sabah emerged as the dominant clan, and its members were formally established as rulers in 1756. By the late 19th century, however, fears of growing Ottoman influence led Sheikh Mubarak Al Sabah to enter into an agreement with Great Britain, which effectively established Kuwait as an autonomous British protectorate. Under the 1899 agreement, Kuwait maintained control over its internal affairs, and Great Britain assumed responsibility for the country's security and foreign relations. By the 1960s, Kuwait enjoyed such a level of economic development that it became one of the wealthiest states on a per capita basis, and in 1961 Kuwait gained full independence from Britain. For a brief period, between 1990 and 1991, Kuwait was under Iraqi control, but the Iraqi army was defeated by United Nation forces.

The official national language of Kuwait is Standard Arabic, a modernized variety of Classical Arabic. The latter, with its archaic vocabulary, is used for religious and ceremonial purposes, whereas Standard Arabic is employed in formal situations, for education, and for official purposes and is the language of the press. In the Arabic-speaking world, only the well educated have adequate proficiency in Standard Arabic; in Kuwait, the literacy rate among men is 85.1% and among women is 81.7%. About 85% of the population uses Gulf Spoken Arabic to communicate. Mehri, a South Arabian language, is used by some 3530 Kuwaitis. In addition, English is widely spoken, and both television and radio broadcasts are often in English. There is one institution in Kuwait for all its 110 285 deaf inhabitants.

See also: Arabic.

Kwa Languages

M E K Dakubu, University of Ghana, Accra, Ghana

The Kwa languages, a branch of the Niger-Congo family, are spoken in West Africa. As presently defined, the group extends from southeastern Ivory Coast in the west through the southern two-thirds of Ghana, southern Togo, and the Republic of Benin to the Benin-Nigeria border in the east. The term Kwa was adopted by linguists toward the end of the 19th century to group together Akan, Ga, Ewe (Éwé), and their close relatives because in many of these languages a stem *kwa* or *kua* means 'man' or 'person.'

Kuryłowicz J (1960/1973). *Esquisses linguistiques.* ('Linguistic sketches'). Wrocław–Warsaw–Cracow: (2nd edn. Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN Munich: Wilhelm Fink).
Kuryłowicz J (1961). *L'Apophonie en semitique* ('Apophony in Semitic'). Wroclaw–Warsaw–Cracow: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.
Kuryłowicz J (1964). *The inflectional categories of Indo–European.* Heidelberg: Carl Winter.
Kuryłowicz J (1968). *Indogermanische Grammatik* ('Indo–European grammar') (vol. 2). Heidelberg: Carl Winter.
Kuryłowicz J (1972). *Studies in Semitic grammar and metrics.* Wroclaw–Warsaw–Cracow: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.
Kuryłowicz J (1975). *Metrik und Sprachgeschichte* ('Metrics and the history of language'). Wroclaw–Warsaw–Cracow–Gdansk: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.
Kuryłowicz J (1977). *Problèmes de linguistique indo-européenne.* ('Problems of Indo–European linguistics'). Wroclaw–Warsaw–Cracow–Gdnask: Zaklad Narodowy im. Ossolinskich & Wydawnictwo PAN.

Kuwait: Language Situation

Editorial Team

On the northwestern coast of the Persian Gulf, Kuwait shares borders with Iraq and Saudi Arabia. Kuwait's modern history began in the early 18th century, when several clans from the Al Aniza tribe migrated to the northern shore of the Gulf from the Najd, their famine-stricken homeland in central Arabia. Eventually, the Al Sabah emerged as the dominant clan, and its members were formally established as rulers in 1756. By the late 19th century, however, fears of growing Ottoman influence led Sheikh Mubarak Al Sabah to enter into an agreement with Great Britain, which effectively established Kuwait as an autonomous British protectorate. Under the 1899 agreement, Kuwait maintained control over its internal affairs, and Great Britain assumed responsibility for the country's security and foreign relations. By the 1960s, Kuwait enjoyed such a level of economic development that it became one of the wealthiest states on a per capita basis, and in 1961 Kuwait gained full independence from Britain. For a brief period, between 1990 and 1991, Kuwait was under Iraqi control, but the Iraqi army was defeated by United Nation forces.

The official national language of Kuwait is Standard Arabic, a modernized variety of Classical Arabic. The latter, with its archaic vocabulary, is used for religious and ceremonial purposes, whereas Standard Arabic is employed in formal situations, for education, and for official purposes and is the language of the press. In the Arabic-speaking world, only the well educated have adequate proficiency in Standard Arabic; in Kuwait, the literacy rate among men is 85.1% and among women is 81.7%. About 85% of the population uses Gulf Spoken Arabic to communicate. Mehri, a South Arabian language, is used by some 3530 Kuwaitis. In addition, English is widely spoken, and both television and radio broadcasts are often in English. There is one institution in Kuwait for all its 110 285 deaf inhabitants.

See also: Arabic.

Kwa Languages

M E K Dakubu, University of Ghana, Accra, Ghana

The Kwa languages, a branch of the Niger-Congo family, are spoken in West Africa. As presently defined, the group extends from southeastern Ivory Coast in the west through the southern two-thirds of Ghana, southern Togo, and the Republic of Benin to the Benin-Nigeria border in the east. The term Kwa was adopted by linguists toward the end of the 19th century to group together Akan, Ga, Ewe (Éwé), and their close relatives because in many of these languages a stem *kwa* or *kua* means 'man' or 'person.'

Constitution of the Group

The list of languages included in Kwa has varied considerably. In the 1950s, Westermann and Bryan included 10 languages of Ivory Coast that they subgrouped into the Lagoon group; Akan, including the languages more recently termed Tano; Ga-Dangme; Ewe; and most of the languages of southern Nigeria. Greenberg's 1963 classification expanded the group to include the Kru languages of Liberia in the east, the languages of the Ghana-Togo hill country that Westermann and Bryan had explicitly excluded, and the Ijaw languages of the Niger delta. This membership was maintained by Stewart in 1971, but in 1989 he eliminated the Kru languages and the Nigerian languages, keeping the Togo Mountain languages and adding a few languages hitherto unknown or unstudied. This is the membership in the most recent overview presented by Williamson and Blench (2000), which essentially has contracted back to the earliest membership plus the Lagoon and the Togo Mountain languages. The group as most recently defined, without Kru and without the Nigerian languages, is sometimes referred to as New Kwa.

Nevertheless, the Kwa languages are currently considered to be more closely related to Yoruba, Igbo, and other languages of southern Nigeria than to the rest of Niger-Congo, forming with them a branch of the Volta-Congo subgroup of Niger-Congo. However, except at the lower levels of classification such as the Tano, Potou-Tano, and Ewe-Fon (Gbe) groups, genetic relationships among these languages are quite distant. It has never been adequately demonstrated using the comparative method that Akan, Ga, Ewe, and the Togo Mountain languages are more closely related to one another than to any other languages.

The Subgroups

The (New) Kwa subgroups are distributed approximately from east to west along the coastal forest and savannah, with the majority of languages and speakers in Ghana. However, many communities have traditions of migration, and it is likely that 1500 years ago their geographical distribution was very different from today. A primary distinction has sometimes been made between the western languages with Ga-Dangme and some of the Togo Mountain languages, referred to as 'Nyo,' and Gbe and the other Togo Mountain languages, referred to as 'Left Bank' (of the Volta River). We examine the composition of the subgroups starting in the west and moving eastward.

The Lagoon languages of the Ivory Coast – Avikam and Alladian, Adjukru (Adioukru), Abidji, Abbey (Abé), Attié, and an isolated language farther west called Ega – are not particularly closely related to one another. Just east of these, the Potou-Tano group comprises at least 20 languages, including Ebrie (Ebrié) and Mbatto (Mbato) (the Potou of Potou-Tano) and the Tano group consisting of Krobu, Abure, and Eotilé (Beti) (all spoken in the Ivory Coast); Nzema, Ahanta, Anyi (Anyin), and Baule (Baoule) (spoken in the eastern Ivory Coast and western Ghana); Anufo (spoken in northeastern Ghana); Akan (with by far the most speakers, mainly in Ghana); and approximately a dozen languages making up the Guang group including Gonja, Nawuri, Nchumuru (Nchumbulu), Krachi (Krache), Nkonya, Okere (Cherepon), Larteh, Efutu, and Awutu (all spoken in Ghana from north of Akan to the sea in the south, distributed along the course of the Volta river). Because the locus of greatest differentiation is in the west, it has been suggested that the Potou-Tano languages spread from west to east. Migration traditions suggest that they also spread in the eastern area (i.e., mainly in Ghana) from north to south.

Ga-Dangme, on the other hand, consists of just two languages, Ga and Dangme (or Adangme). Now situated on the Accra Plain south east of the Tano languages and west of the Volta River, the speakers have traditions of migration from the east, possibly in conjunction with the Ewe, and from the northeast, probably from what is now the northern Republic of Benin.

The Central Togo or Togo Mountain languages comprise 14 languages spoken in Ghana, Togo, and the Republic of Benin, most of them in the hilly area on the Ghana-Togo border. An earlier name for these languages was Togo Remnant Languages, adapted from the German *Togorestsprachen*. That name reflected the possibility that their fragmented distribution on hilltops difficult to access and their generally small numbers of speakers result from these linguistic communities having fled to refuge areas at some time in the past, under pressure from expanding larger groups such as (perhaps) the Ewe and the Akan. True or not, the name is disliked by speakers and no longer used.

Heine (1968) treated the languages as a genetic family and classified them into two subgroups: the NA subgroup (Basila (Anii), Lelemi-Lefana (Lelemi), Logba, Adele, Likpe (Sekpele), Santrokofi (Sele), and Akpafu-Lolobi (Siwu)) and the KA subgroup (Avatime, Nyangbo-Tafi, Bowiri (Tuwili), Ahlo, Kposo (Akposo), Kebu (Akebou), and Animere). More recent classifications accept these two groupings, but

put them directly under Kwa; that is, they are now considered no more closely related to one another than to the other New Kwa groups, and 'Togo Mountain' or 'Central Togo languages' are merely geographical labels.

Kposo has the most speakers by a considerable margin, approximately 80 000, followed by Kebu with 17 500. (Both are spoken mainly in Togo; figures are based on Heine, 1968.) Lelemi-Lefana has approximately 15 000 speakers and Adele approximately 8000. Others have fewer: Logba reportedly has approximately 2000 speakers, and Animere was said to have fewer than 300 in the 1960s. There is evidence that one or two languages have died out in the area in the course of the past 2 centuries.

The area extending approximately 80 miles inland along the coast from the Volta eastward to the Nigerian border is dominated by the closely knit Gbe group of languages and dialects. The name is adopted from their common word for 'language'. They include Ewe, spoken in Ghana and Togo by more than 3 million people. The Ewe-speaking people have a tradition of migration from farther east but still within the Gbe area, from Nuatja and Tado in southern Togo. Gen (Gen-gbe) is the language of Lomé, the capital of Togo and a major lingua franca. Another major variety is Aja (Aja-gbe). Somewhat more distantly related are Maxi (Maxi-gbe) and Fon (Fon-gbe), the language of the old kingdom of Dahomey, which is now part of the Republic of Benin. The easternmost variety is Gun (Gun-gbe), which crosses the Nigerian border along the coast and is thought to have been spoken in the Lagos area in earlier times.

Linguistic Characteristics

In all Kwa languages, the fundamental word order is subject-verb-object. It was once thought that they typically had very simple morphology, but, although the situation varies among the different groups, this is generally not really the case. Part of the reason for this belief may have been the fact that in many of the languages tone plays an important role in morphology, with some grammatical morphemes being realized only by tone. Thus, in the following Ga sentence only the high tone on the second syllable of the pronoun *amɛ* 'they' shows that the verb is perfect and not aorist.

amɛ́-bá
'they have come'
they.PERF-come

Tone is also important in making lexical distinctions; again in Ga, we have *lá* 'blood' with high tone but *là* 'fire' with low. The number of phonemic tones ranges from the minimal two in Ga and Akan to five in Avatime. In all the languages west of the Volta, tone levels are affected by downstepping, or the lowering of all pitch levels after a low tone throughout a sentence, but many languages east of the Volta, including Ewe, do not have this.

Cross-height vowel harmony, based on the position of the tongue root, is typical of the Tano languages but not of Ga-Dangme, Gbe, or some of the Togo Mountain languages. In all (Adjukru is an exception), syllables are typically open, ending in a vowel or sometimes a syllabic consonant.

Double articulated consonants ([*k͡p*], less often [*g͡b*]) are typical of Ga-Dangme, the Togo Mountain languages, and Gbe but not of Akan and some other Potou-Tano languages. In Ga-Dangme and Gbe (and also Yoruba), but again not in the western languages, the voiceless bilabial stop /p/ did not exist until it was introduced through loanwords in the course of the 18th century. Consonant clusters either do not exist or the second consonant is limited to off-glides and /l/ or /r/ and the cluster is analyzable as being derived from a CVV or CVCV structure.

Verb systems vary, but paradigms generally express aspect features, with the perfective-imperfective distinction basic. Grammatical tense is less important, often limited to a contrast between future and nonfuture, and even that shows signs of having been grammaticized relatively recently in many languages. The expression of progressive aspect is of particular interest because the western languages differ radically in its expression from many of those spoken east of the Volta. In Tano languages and also in Ga, the progressive (like most other aspect features) is shown by a prefix, as in the following sentences in Akan (1) and Ga (2).

(1) abofra no re-di nneɛma
child the PROG-eat things
'the child is eating something'

(2) gbekɛ lɛ mii-ye nii
child the PROG-eat things
'the child is eating something'

In Ewe (3) and Dangme (4), on the other hand, it is periphrastically expressed, with an aspect verb followed by a nonfinite form of the event verb preceded by its object.

(3) Ɖevi a le nu ɖu-m
child the is thing eat-PROG
'the child is eating (something)'

(4) jukwɛ ɔ ŋɛ nɔ ye-e
child the is thing eat-PROG
'the child is eating (something)'

Some of the Togo Mountain languages use the periphrastic style of expression, but many do not.

In Kwa languages, the noun is followed by the adjective, article, and so on that modify it, whereas a possessor precedes the possessed noun. However, there are radical differences in nominal morphology. The Togo Mountain languages generally have several singular and plural classes expressed by prefixes; modifiers take prefixes that show agreement (concord) with the class of the noun they modify and pronouns also vary according to the class of the noun referred to. Akan and other Tano languages have several singular and plural prefixes for nouns but no class concord. Ga-Dangme, however, uses basically just one plural (no singular) suffix, whereas the Gbe languages use a phrase-final particle, as in the following example from Fon.

(5) xɔ̀ntɔ̀n yétɔ̀n wè lɛ́
friend your 2 PL
'both your friends'

Serial verb constructions, in which several verbs share one subject and sometimes object with no intervening conjunction, occur in all Kwa languages.

See also: Akan; Benin: Language Situation; Class Language; Ewe; Ghana: Language Situation; Ivory Coast: Language Situation; Migration and Language; Niger-Congo Languages; Nigeria: Language Situation; Phonetics of Harmony Systems; Tense, Mood, Aspect: Overview; Togo: Language Situation; Tone: Phonology.

Bibliography

Capo H C (1991). *A comparative phonology of Gbe.* Berlin: Mouton de Gruyter.

Dakubu M E K (ed.) (1988). *The languages of Ghana.* London: Kegan Paul International.

Dakubu M E K & Osam E K (eds.) (2004). *Studies in the languages of the Volta Basin* (vol. 2). Legon, Ghana: University of Ghana Linguistics Department.

Heine B (1968). *Die Verbreitung und Gliederung der Togorestsprachen.* Berlin: Dietrich Reimer Verlag.

Herault G (ed.) (1982). *Atlas des langues kwa de Côte d'Ivoire.* Abidjan, Côte d'Ivoire: Université d'Abidjan Institut de Linguistique Appliquée.

Leben W R & Ahoua F (eds.) (2002). *Wortkunst und Dokumentartexte in Afrikanischen Sprachen 12: Contes et textes documentaires kwa de Côte d'Ivoire.* Cologne: Rüdiger Köppe.

Stewart J M (1971). 'Niger-Congo, Kwa.' In Sebeok T A (ed.) *Current trends in linguistics, vol. 7: Linguistics in sub-Saharan Africa.* The Hague and Paris: Mouton. 179–212.

Stewart J M (1989). 'Kwa.' In Bendor-Samuel J (ed.) *The Niger-Congo languages.* New York: University Press of America. 217–245.

Westermann D & Bryan M A (1952). *Handbook of African languages part II: Languages of West Africa.* London: Oxford University Press.

Williamson K & Blench R (2000). 'Niger-Congo.' In Heine B & Nurse D (eds.) *African languages, an introduction.* Cambridge, UK: Cambridge University Press. 11–42.

Kyrgyzstan: Language Situation

L Johanson, Mainz University, Mainz, Germany

Languages

Kyrgyzstan is a Central Asian country (capital: Bishkek, until 1990 named Frunze), bordering on China, Uzbekistan, Tajikistan, and Kazakhstan. It was annexed by Russia in the late 19th century, became a Soviet Socialist Republic in 1936, and achieved independence from the Soviet Union in 1991. Its official languages are Kirghiz and Russian.

The population is mainly concentrated to the lowlands of the Chu and Talas Valleys and the Kirghiz part of the Ferghana Valley. According to the 1999 census, the main linguistic groups among the roughly 5 million inhabitants are Kirghiz (64.9%), Uzbeks (13.8%), and Russians (12.5%). The rest is composed of speakers of numerous other languages. The seemingly weak position of the Uyghurs, who make up 1% of the population, is partly due to their tendency to refer to themselves as Uzbek.

The Ferghana Valley, located on Upper Syrdarya, is divided between Kyrgyzstan, Uzbekistan, and Tajikistan. As a result of the administrative division set up in the Stalin era, the borders of the three countries do not follow the borders of the speech communities. The valley has remained a matter of dispute between the countries. The part of the valley that belongs to Kyrgyzstan, even its main city Osh, has a predominantly Uzbek population. Many speakers of Kirghiz live on Uzbek and Tajik territories. Two enclaves in Uzbekistan and one in Tajikistan belong to Kyrgyzstan, whereas several small Uzbek enclaves exist in Kyrgyzstan. In 1992–1993, refugees from Tajikistan moved into southern Kyrgyzstan.

Some of the Togo Mountain languages use the periphrastic style of expression, but many do not.

In Kwa languages, the noun is followed by the adjective, article, and so on that modify it, whereas a possessor precedes the possessed noun. However, there are radical differences in nominal morphology. The Togo Mountain languages generally have several singular and plural classes expressed by prefixes; modifiers take prefixes that show agreement (concord) with the class of the noun they modify and pronouns also vary according to the class of the noun referred to. Akan and other Tano languages have several singular and plural prefixes for nouns but no class concord. Ga-Dangme, however, uses basically just one plural (no singular) suffix, whereas the Gbe languages use a phrase-final particle, as in the following example from Fon.

(5)	xɔ̀ntɔ̀n	yétɔ̀n	wè	lɛ́
	friend	your	2	PL
	'both your friends'			

Serial verb constructions, in which several verbs share one subject and sometimes object with no intervening conjunction, occur in all Kwa languages.

See also: Akan; Benin: Language Situation; Class Language; Ewe; Ghana: Language Situation; Ivory Coast: Language Situation; Migration and Language; Niger-Congo Languages; Nigeria: Language Situation; Phonetics of Harmony Systems; Tense, Mood, Aspect: Overview; Togo: Language Situation; Tone: Phonology.

Bibliography

Capo H C (1991). *A comparative phonology of Gbe.* Berlin: Mouton de Gruyter.

Dakubu M E K (ed.) (1988). *The languages of Ghana.* London: Kegan Paul International.

Dakubu M E K & Osam E K (eds.) (2004). *Studies in the languages of the Volta Basin* (vol. 2). Legon, Ghana: University of Ghana Linguistics Department.

Heine B (1968). *Die Verbreitung und Gliederung der Togorestsprachen.* Berlin: Dietrich Reimer Verlag.

Herault G (ed.) (1982). *Atlas des langues kwa de Côte d'Ivoire.* Abidjan, Côte d'Ivoire: Université d'Abidjan Institut de Linguistique Appliquée.

Leben W R & Ahoua F (eds.) (2002). *Wortkunst und Dokumentartexte in Afrikanischen Sprachen 12: Contes et textes documentaires kwa de Côte d'Ivoire.* Cologne: Rüdiger Köppe.

Stewart J M (1971). 'Niger-Congo, Kwa.' In Sebeok T A (ed.) *Current trends in linguistics, vol. 7: Linguistics in sub-Saharan Africa.* The Hague and Paris: Mouton. 179–212.

Stewart J M (1989). 'Kwa.' In Bendor-Samuel J (ed.) *The Niger-Congo languages.* New York: University Press of America. 217–245.

Westermann D & Bryan M A (1952). *Handbook of African languages part II: Languages of West Africa.* London: Oxford University Press.

Williamson K & Blench R (2000). 'Niger-Congo.' In Heine B & Nurse D (eds.) *African languages, an introduction.* Cambridge, UK: Cambridge University Press. 11–42.

Kyrgyzstan: Language Situation

L Johanson, Mainz University, Mainz, Germany

Languages

Kyrgyzstan is a Central Asian country (capital: Bishkek, until 1990 named Frunze), bordering on China, Uzbekistan, Tajikistan, and Kazakhstan. It was annexed by Russia in the late 19th century, became a Soviet Socialist Republic in 1936, and achieved independence from the Soviet Union in 1991. Its official languages are Kirghiz and Russian.

The population is mainly concentrated to the lowlands of the Chu and Talas Valleys and the Kirghiz part of the Ferghana Valley. According to the 1999 census, the main linguistic groups among the roughly 5 million inhabitants are Kirghiz (64.9%), Uzbeks (13.8%), and Russians (12.5%). The rest is composed of speakers of numerous other languages. The seemingly weak position of the Uyghurs, who make up 1% of the population, is partly due to their tendency to refer to themselves as Uzbek.

The Ferghana Valley, located on Upper Syrdarya, is divided between Kyrgyzstan, Uzbekistan, and Tajikistan. As a result of the administrative division set up in the Stalin era, the borders of the three countries do not follow the borders of the speech communities. The valley has remained a matter of dispute between the countries. The part of the valley that belongs to Kyrgyzstan, even its main city Osh, has a predominantly Uzbek population. Many speakers of Kirghiz live on Uzbek and Tajik territories. Two enclaves in Uzbekistan and one in Tajikistan belong to Kyrgyzstan, whereas several small Uzbek enclaves exist in Kyrgyzstan. In 1992–1993, refugees from Tajikistan moved into southern Kyrgyzstan.

Language Policy

Kirghiz was established as a standard language in the Soviet period. A written variety was created in 1923. The originally used Arabic script was abandoned for a Roman-based alphabet; a Cyrillic alphabet was adopted in 1940–1941.

In the early 1990s, Kirghiz was introduced as the official language of the independent republic. In 2001, after substantial pressure from Russia, the Russian language became the second official state language alongside Kirghiz.

In the first years of independence, a significant number of non-Kirghiz speakers left the republic. The emigration rate was highest in 1993, but decreased subsequently thanks to measures against discrimination against citizens who did not master Kirghiz. The number of Russians decreased by at least 6%. Practically all speakers of German, who had made up 1% of the population, left the country. On the other hand, the Kirghiz-speaking proportion of the population increased due to immigration from Tajikistan and Uzbekistan.

Though Russian still dominates public life in Kyrgyzstan, e.g., by being the main medium of publication (newspapers, books, etc.), the efforts to protect and promote Kirghiz have been successful. The Kirghiz generally maintain a high degree of mastery of their language. A return to a Roman-based alphabet has been discussed and planned, but not implemented so far.

See also: Kirghiz; Russian; Tajikistan: Language Situation; Uzbek; Uzbekistan: Language Situation.

Language Maps (Appendix 1): Map 107.

Bibliography

Imart G (1981). *Le kirghiz. 1–2.* Aix-en-Provence: Publications de l'Université de Provence.

Johanson L & Csató É Á (eds.) (1998). *The Turkic languages.* London: Routledge.

Landau J M & Kellner-Heinkele B (2001). *Politics of language in the ex-Soviet Muslim states: Azerbaijan, Uzbekistan, Kazakhstan, Kyrgyzstan, Turkmenistan, Tajikistan.* Ann Arbor: University of Michigan Press.

Relevant Website

http://www.turkiclanguages.com.

Langdon M (1986). 'Hokan-Siouan revisited.' In Cowan W, Foster M K & Koerner K (eds.) *New perspectives in language, culture, and personality.* Amsterdam: Benjamins. 111–146.
Langdon M & Silver S (eds.) (1976). *Hokan studies.* The Hague: Mouton.
Langdon M & Jacobsen W (1996). 'Selected bibliography of recent work relevant to Hokan.' *SCOIL 9,* 131–132.
Leŝĉiner D R (1989). 'Rekonstrukcija nazvanij ĉastej tela v jazykax sem'i xoka.' In *Lingvistiĉeskaja Rekonstrukcija i Drevnejŝaja Istorija Vostoka: Materialy k diskussijam na Meẑdunarodnoj konferencii (Moskva, 29 maja-ijunja 1989g.).* Ĉast' 1. Moscow: Izdatel'stvo Nauka. 159–166.
McLendon S (1973). *University of California Publications in Linguistics 71: Proto-Pomo.* Berkeley: University of California Press.
Mithun M (1999). *The languages of Native North America.* Cambridge: Cambridge University Press.
Oswalt R L (1977). 'The Hokan word for "tongue": the Pomo evidence.' *University Museum Studies (Southern Illinois University) 11,* 69–78.
Powell J W (1891). 'Indian linguistic families of America north of Mexico.' *Bureau of American Ethnology Annual Report 7,* 1–142.
Sapir E (1917). *University of Caliifornia Publications in American Archeology and Ethnology 13*: The position of Yana in the Hokan stock.
Sapir E (1920a). 'The Hokan and Coahuiltecan languages.' *IJAL 1,* 280–290.
Sapir E (1920b). 'A note on the first person plural in Chimariko.' *International Journal of American Linguistics 1,* 291–294.
Sapir E (1921). 'A supplementary note on Salinan and Washo.' *International Journal of American Linguistics 2,* 68–72.
Sapir E (1925). 'The Hokan affinity of Subtiaba in Nicaragua.' *American Anthropologist 27,* 402–435, 491–527.
Swanton J R (1915). 'Linguistic position of the tribes of Southern Texas and Northeastern Mexico.' *American Anthropologist 17,* 17–40.
Wares A C (1968). *A comparative study of Yuman consonantism.* The Hague: Mouton.

Holcot, Robert (d. 1349)

E Bell Canon, Social Circle, GA, USA

Robert Holcot (Holkot) was born near Northampton, England c. 1290. Joining the Dominican order, he arrived at Oxford and would remain there as student and then teacher from 1326–1332. Holcot may have served in the capacity of Dominican regent master at Oxford until being called to the post of clerk for Richard de Bury, Bishop of Durham. In 1343, Holcot returned to Northampton and lived in the priory there until his death of the bubonic plague in 1349.

Dominican training at the time of Robert Holcot's education would have likely included the study of logic, grammar, philosophy, and theology. His training on the differences between faith and reason would have followed the teaching of Saint Thomas Aquinas. Holcot himself was more a follower of William of Ockham than of Aquinas. In fact, although not fully agreeing with Ockham on matters of epistemology with regards to intuition and abstraction, Holcot used other fundamental principles from the teachings and writings of Ockham as a basis for his religious philosophy.

Holcot tried to reconcile the impassable differences between logic and faith. Since the principle of noncontradiction is fundamental to a logical progression of thought, it is difficult to reconcile the Christian belief that God is omnipotent within the framework of that logical progression. For instance, 'If God is omnipotent, can God create a rock so big that even God can't move it?' For Holcot, the answer was that natural laws with regard to logic and the limitations of our own human language must not apply to God. There must be, in addition to Aristotelian logic, which is applicable to the natural order, a separate logic of faith that is applicable to God.

Also problematic for the faithful Christian was the notion of future contingents with regard to divine omniscience. In other words, if God knows all things and any contingency is able to either exist or not, God cannot have foreknowledge of something that is inherently contingent. Holcot reconciled this disagreement between the logical and the theological by arguing that (simply put) if God is omniscient, the debate over whether or not he can know the nature of a future contingency is moot.

The written works of Robert Holcot include his comments on Peter Lombard's *Sentences*. Written while at Oxford, Holcot's 'Comments' on the *Sentences* are considered to be his most famous work. His *Sermo finalis*, given at the end of his tenure at Oxford, employs double entendre, or puns, to address his disapproval of his successor. Also of importance, the *Sex articuli* is both a comment on epistemology and a criticism of Holcot's contemporaries. It is possible that the *Sex articuli* was intended to appear together with his comments on the *Sentences*. Other important works include comments on Matthew,

De imputabilitate peccati (On the imputability of sin); *De stellis* (On the Stars), *Moralitates* and *Convertimini; Postilla super librum Sapientiae;* and commentaries on Proverbs, Canticle of Canticles, and Ecclesiasticus.

Bibliography

Gelber H 'Robert Holkot.' *The Stanford encyclopedia of philosophy (Fall 2001 edition)*, Zalta E N (ed.). http://plato.stanford.edu/archives/fall2001/entries/holkot/.

Holcot R (1995). *Seeing the future clearly, questions of future contingents by Robert Holcot*. Streveler P A & Tachau K H *et al.* (ed.). Toronto: Pontifical Institute of Mediaeval Studies.

Maurer A (1967). 'Robert Holcot.' In *New Catholic encyclopedia*, 19 vols. The Catholic University of America. New York: McGraw-Hill, 1967.

Holder, William (1616–1698)

C Neis, Universitat Potsdam, Potsdam, Germany

William Holder was born in 1616 in Southwell, Nottinghamshire, England and died on January 24, 1698 in Hertford, England. Holder was a clergyman, mathematician, musician, phonetician, and speech therapist. He was the first Englishman in the 17th century to propose an explicit practical application of an advanced phonetic theory.

Holder began his studies at Pembroke Hall, Cambridge in 1633. He was elected Fellow in 1640, after his graduation. During the Civil War, he moved to Oxford, where he became a doctor of divinity at Oxford University in 1660. In 1663 he was elected to the Royal Society, which recognized his achievements in phonetics and his system of teaching the deaf and mute people. Apart from his phonetic interests, Holder also made important contributions to music theory, especially to the theory of harmony and composition.

From 1659 to 1660, he taught 10-year-old Alexander Popham, who was both deaf and mute, to speak. Holder described this experience in *Philosophical Transactions of the Royal Society* (1668), but his successful language training to Popham soon caused a public quarrel between Holder and John Wallis. Wallis, a mathematician who was a well-known instructor of deaf-mutes, taught Popham after the boy had worked with Holder. Wallis, however, credited himself alone with the successful training of Popham without mentioning Holder's merits. Holder gave his account of the dispute in his *A supplement to the philosophical transactions of July 1670, with some reflexions on Dr. John Wallis, his letter there inserted* (1678).

Holder's work with Popham had a deep impact on his phonetic theory, which he explained in his *Elements of speech: an essay of inquiry into the natural production of letters: with an appendix concerning persons deaf and dumb* (1669). The appendix was a reprint of Holder's report on his work with Popham in *Philosophical transactions*; it may be regarded as the first definitive manual of speech therapy written in England. *Elements* was a particularly valuable piece of English linguistic scholarship in the 17th century, because Holder was the first Englishman to spell out an explicit application of a thoroughly developed phonetic theory. By applying an empirical method based on inductive principles, he sought to analyze the speech act itself. Having verified his experiments, he tried to proceed to generalizations. In *Elements*, Holder proposed a classification of human sounds by referring to the manner and place of articulation. For this purpose, he developed a system of contrastive analysis. His treatise shows that he was also aware of suprasegmental phenomena, e.g., 'accent' and 'emphasis.'

In the *Appendix*, he ascribed the muteness of deaf people to hearing impairment and not to articulatory deficits. His method of teaching was similar to that of Johann Conrad Amman; both relied on the visual and haptic experience of their students as they showed them the respective position of the articulatory organs in regard to each sound. The students were also encouraged to touch their teacher's articulators. It was mainly Holder's belief in the compensatory capacities of his students' senses that inspired him to use this method.

Holder was not only a speech therapist, but also a clergyman and musician. A canon at St Paul's cathedral in 1672, he was appointed sub-dean of the Chapel Royal in 1674. Holder married the sister of Sir Christopher Wren and heavily influenced Wren's education. Both Holder and his wife were buried in St Paul's.

De imputabilitate peccati (On the imputability of sin); *De stellis* (On the Stars), *Moralitates* and *Convertimini; Postilla super librum Sapientiae;* and commentaries on Proverbs, Canticle of Canticles, and Ecclesiasticus.

Bibliography

Gelber H 'Robert Holkot.' *The Stanford encyclopedia of philosophy (Fall 2001 edition)*, Zalta E N (ed.). http://plato.stanford.edu/archives/fall2001/entries/holkot/.

Holcot R (1995). *Seeing the future clearly, questions of future contingents by Robert Holcot.* Streveler P A & Tachau K H *et al.* (ed.). Toronto: Pontifical Institute of Mediaeval Studies.

Maurer A (1967). 'Robert Holcot.' In *New Catholic encyclopedia,* 19 vols. The Catholic University of America. New York: McGraw-Hill, 1967.

Holder, William (1616–1698)

C Neis, Universitat Potsdam, Potsdam, Germany

William Holder was born in 1616 in Southwell, Nottinghamshire, England and died on January 24, 1698 in Hertford, England. Holder was a clergyman, mathematician, musician, phonetician, and speech therapist. He was the first Englishman in the 17th century to propose an explicit practical application of an advanced phonetic theory.

Holder began his studies at Pembroke Hall, Cambridge in 1633. He was elected Fellow in 1640, after his graduation. During the Civil War, he moved to Oxford, where he became a doctor of divinity at Oxford University in 1660. In 1663 he was elected to the Royal Society, which recognized his achievements in phonetics and his system of teaching the deaf and mute people. Apart from his phonetic interests, Holder also made important contributions to music theory, especially to the theory of harmony and composition.

From 1659 to 1660, he taught 10-year-old Alexander Popham, who was both deaf and mute, to speak. Holder described this experience in *Philosophical Transactions of the Royal Society* (1668), but his successful language training to Popham soon caused a public quarrel between Holder and John Wallis. Wallis, a mathematician who was a well-known instructor of deaf-mutes, taught Popham after the boy had worked with Holder. Wallis, however, credited himself alone with the successful training of Popham without mentioning Holder's merits. Holder gave his account of the dispute in his *A supplement to the philosophical transactions of July 1670, with some reflexions on Dr. John Wallis, his letter there inserted* (1678).

Holder's work with Popham had a deep impact on his phonetic theory, which he explained in his *Elements of speech: an essay of inquiry into the natural production of letters: with an appendix concerning persons deaf and dumb* (1669). The appendix was a reprint of Holder's report on his work with Popham in *Philosophical transactions*; it may be regarded as the first definitive manual of speech therapy written in England. *Elements* was a particularly valuable piece of English linguistic scholarship in the 17th century, because Holder was the first Englishman to spell out an explicit application of a thoroughly developed phonetic theory. By applying an empirical method based on inductive principles, he sought to analyze the speech act itself. Having verified his experiments, he tried to proceed to generalizations. In *Elements*, Holder proposed a classification of human sounds by referring to the manner and place of articulation. For this purpose, he developed a system of contrastive analysis. His treatise shows that he was also aware of suprasegmental phenomena, e.g., 'accent' and 'emphasis.'

In the *Appendix*, he ascribed the muteness of deaf people to hearing impairment and not to articulatory deficits. His method of teaching was similar to that of Johann Conrad Amman; both relied on the visual and haptic experience of their students as they showed them the respective position of the articulatory organs in regard to each sound. The students were also encouraged to touch their teacher's articulators. It was mainly Holder's belief in the compensatory capacities of his students' senses that inspired him to use this method.

Holder was not only a speech therapist, but also a clergyman and musician. A canon at St Paul's cathedral in 1672, he was appointed sub-dean of the Chapel Royal in 1674. Holder married the sister of Sir Christopher Wren and heavily influenced Wren's education. Both Holder and his wife were buried in St Paul's.

Holder owes his reputation as a phonetician and speech teacher before all to his *Elements*, which may be regarded as a milestone in the establishment of phonetics as a science.

See also: Amman, Johann Conrad (1669–1730); Phonetics: Overview; Sign Language: Communities and Cultures; Wallis, John (1616–1703).

Bibliography

Abercrombie D (1965). 'Forgotten phoneticians.' *Studies in Phonetics and Linguistics* (Language and language learning; 10), 44–75.

Gessinger J (1994). *Auge & Ohr. Studien zur Erforschung der Sprache am Menschen.* Berlin/New York: de Gruyter.

Hartley Sir H (ed.) (1960). *The Royal Society: its origins and founders.* London: The Royal Society London.

Holder W (1668). 'An experiment concerning deafness.' *Philosophical Transactions of the Royal Society,* 665–668.

Holder W (1669). *Elements of speech: an essay of inquiry into the natural production of letters: with an appendix concerning persons deaf and dumb.* London. Printed by T. N. for J. Martyn, Printer to the R. Society at the Bell without Temple-Barr.

Holder William (1669/1967). *Elements of speech. (English linguistics, 1500–1800).* Menston: Scolar Press. [Reprint.]

Holder W (1678). *A supplement to the philosophical transactions of July 1670, with some reflexions on Dr. John Wallis, his letter there inserted.* London: Henry Brome.

Holder W (1694/1967). *Treatise of the natural grounds and principals of harmony.* (*Monuments of music and music literature in facsimile.* Ser. 2.32). New York: Broude Brothers.

Purver M (1967). *The Royal Society: concept and creation.* With an introduction by H R Trevor-Roper. London: Routledge and Kegan Paul.

Salmon V (1979). *The study of language in 17th-century England. (Studies in the History of Linguistics, vol. 17).* Amsterdam: Benjamins.

Subbiondo J L (1978). 'William Holder's *Elements of speech (1969):* a study of applied English phonetics and speech therapy.' *Lingua 46,* 169–184.

Subbiondo J L (1996). 'Holder, William.' In Stammerjohann H (ed.) *Lexicon grammaticorum. Who's who in the history of world linguistics.* Tübingen: Niemeyer. 424–426.

Tilmouth M & Stanley J (2002). 'Holder William.' In Sadie S (ed.) *The new grove dictionary of music and musicians,* (2nd edn., vol. 11, Harpégé to Hutton). 618.

Holism, Semantic and Epistemic

R Vallée, Université de Moncton, Shippagan, Canada

According to Epistemic Holism, one does not test hypotheses – individual sentences – one at a time; rather the process of testing hypotheses confronts the whole of science with reality. According to Semantic Holism, the meaning of words and sentences in a language depends on the meaning of each and every word/sentence in that language. The former is widely accepted, whereas the latter is very controversial. Both forms of holism are tied together by the idea that the meaning of sentences is given by their truth conditions.

Epistemic Holism

Scientific theories confront reality via hypotheses, and *prima facie*, hypotheses are tested one at the time. Carl Gustav Hempel (1905–1997) and Willard van Orman Quine (1908–2000) contended that this view was an oversimplification and argued for what is now known as the Duhem/Quine thesis or Epistemic Holism (EH): "Science faces the tribunal of experience not sentence by sentence but as a corporate body: the whole of science" (Quine, 1986: 620).

Science is not compartmentalized. For example, observations made by telescopes are backed by astronomy, physics, optics, and the mathematical theories used in all these theories. The argument behind EH is simple. Assume a logical empiricist's theory of meaning according to which the meaning of a sentence is given by its confirmation or verification conditions. Consider now any theoretical sentence. Such a sentence cannot be reduced to one single sentence concerning experience and connecting it directly to the world. So, it has no confirmation conditions as a single, independent item. It is connected to reality through a net of sentences, some of which get their truth conditions straight from facts with no regard for other sentences, the so-called observation sentences. It follows that a hypothesis is just a sensor of the whole science, which confronts reality and is affected as a whole through the hypothesis. Hempel (1950) wrote:

> "In order to understand 'the meaning' of a hypothesis within an empiricist language, we have to know not merely what observation sentences it entails alone or in conjunction with subsidiary hypothesis, but also what

Holder owes his reputation as a phonetician and speech teacher before all to his *Elements*, which may be regarded as a milestone in the establishment of phonetics as a science.

See also: Amman, Johann Conrad (1669–1730); Phonetics: Overview; Sign Language: Communities and Cultures; Wallis, John (1616–1703).

Bibliography

Abercrombie D (1965). 'Forgotten phoneticians.' *Studies in Phonetics and Linguistics* (Language and language learning; 10), 44–75.

Gessinger J (1994). *Auge & Ohr. Studien zur Erforschung der Sprache am Menschen.* Berlin/New York: de Gruyter.

Hartley Sir H (ed.) (1960). *The Royal Society: its origins and founders.* London: The Royal Society London.

Holder W (1668). 'An experiment concerning deafness.' *Philosophical Transactions of the Royal Society,* 665–668.

Holder W (1669). *Elements of speech: an essay of inquiry into the natural production of letters: with an appendix concerning persons deaf and dumb.* London. Printed by T. N. for J. Martyn, Printer to the R. Society at the Bell without Temple-Barr.

Holder William (1669/1967). *Elements of speech. (English linguistics, 1500–1800).* Menston: Scolar Press. [Reprint.]

Holder W (1678). *A supplement to the philosophical transactions of July 1670, with some reflexions on Dr. John Wallis, his letter there inserted.* London: Henry Brome.

Holder W (1694/1967). *Treatise of the natural grounds and principals of harmony.* (*Monuments of music and music literature in facsimile.* Ser. 2.32). New York: Broude Brothers.

Purver M (1967). *The Royal Society: concept and creation.* With an introduction by H R Trevor-Roper. London: Routledge and Kegan Paul.

Salmon V (1979). *The study of language in 17th-century England. (Studies in the History of Linguistics, vol. 17).* Amsterdam: Benjamins.

Subbiondo J L (1978). 'William Holder's *Elements of speech (1969):* a study of applied English phonetics and speech therapy.' *Lingua 46,* 169–184.

Subbiondo J L (1996). 'Holder, William.' In Stammerjohann H (ed.) *Lexicon grammaticorum. Who's who in the history of world linguistics.* Tübingen: Niemeyer. 424–426.

Tilmouth M & Stanley J (2002). 'Holder William.' In Sadie S (ed.) *The new grove dictionary of music and musicians,* (2nd edn., vol. 11, Harpégé to Hutton). 618.

Holism, Semantic and Epistemic

R Vallée, Université de Moncton, Shippagan, Canada

According to Epistemic Holism, one does not test hypotheses – individual sentences – one at a time; rather the process of testing hypotheses confronts the whole of science with reality. According to Semantic Holism, the meaning of words and sentences in a language depends on the meaning of each and every word/sentence in that language. The former is widely accepted, whereas the latter is very controversial. Both forms of holism are tied together by the idea that the meaning of sentences is given by their truth conditions.

Epistemic Holism

Scientific theories confront reality via hypotheses, and *prima facie*, hypotheses are tested one at the time. Carl Gustav Hempel (1905–1997) and Willard van Orman Quine (1908–2000) contended that this view was an oversimplification and argued for what is now known as the Duhem/Quine thesis or Epistemic Holism (EH): "Science faces the tribunal of experience not sentence by sentence but as a corporate body: the whole of science" (Quine, 1986: 620).

Science is not compartmentalized. For example, observations made by telescopes are backed by astronomy, physics, optics, and the mathematical theories used in all these theories. The argument behind EH is simple. Assume a logical empiricist's theory of meaning according to which the meaning of a sentence is given by its confirmation or verification conditions. Consider now any theoretical sentence. Such a sentence cannot be reduced to one single sentence concerning experience and connecting it directly to the world. So, it has no confirmation conditions as a single, independent item. It is connected to reality through a net of sentences, some of which get their truth conditions straight from facts with no regard for other sentences, the so-called observation sentences. It follows that a hypothesis is just a sensor of the whole science, which confronts reality and is affected as a whole through the hypothesis. Hempel (1950) wrote:

> "In order to understand 'the meaning' of a hypothesis within an empiricist language, we have to know not merely what observation sentences it entails alone or in conjunction with subsidiary hypothesis, but also what

> other, non observational, empirical sentences are entailed by it, and for what other hypotheses the given one would be confirmatory or disconfirmatory. In other words, the cognitive meaning of a statement in an empirical language is reflected in the totality of its logical relationships to all other statements in that language, and not to the observation sentences alone."

In that conception, there is no clear-cut distinction between changing the theory and changing the meaning of sentences. When a hypothesis is proven false, one can alter either other parts of the theory, the so-called auxiliary hypothesis, or the meaning of the sentences of the theory. EH is sometimes offered as a reason for Semantic Holism: EH and the empiricist theory of meaning lead to SH. If the meaning of a sentence is determined by its verification conditions and if they are holistic, the meaning of a sentence is also determined holistically. EH and Semantic Holism are independent views on different topics: scientific theories and language. However, they dovetail.

Semantic Holism

Although EH is sentence oriented, Semantic Holism (SH) covers both sentential and subsentential items. There are different versions of SH, but its basic tenet is that the meaning of a sentence/word is determined by, depends on, or is constituted by, its relation to *all* the sentences/words in a language.

The Argument from Compositionality and the Context Principle

SH is usually not advocated in and of itself. EH can motivate it. It can also be introduced as a consequence of two independently motivated, nonepistemic linguistic principles and a truth-based semantics. It is based on two principles:

1. *Principle of Compositionality*: The meaning of a sentence is a function of the meaning of its constituents and their syntactic relationship
2. *Context Principle:* "Only in the context of a sentence does a word have meaning" (Frege, 1978: x).

To give the meaning of a word, one must give the meaning of all sentences in which it occurs; that is, the truth conditions of all these sentences. Davidson (1967: 22) followed the argument to its obvious conclusion that "only in the context of a language does a sentence (and therefore a word) have meaning." The meaning of a word or a sentence is determined by the language to which it belongs.

SH opposes 'Semantic Atomism,' which holds that the meaning of a sentence/word is *independent* from the meaning of *all* the other sentences/words in a language. Semantic Atomism proposes either the referent of a word or an abstract entity, like a Fregean sense, as the meaning of a word. From then on, meanings can be piled up.

SH also opposes '*Semantic Molecularism*,' which holds that the meaning of a sentence/word is determined by, depends on, or is constituted by its relation to *some* of the sentences/words in a language.

Both the Principle of Compositionality (PC) and Context Principle (CP) are *a priori*, methodological principles in theoretical semantics. Semantic Holism *per se* is not methodological in nature. Following from two *a priori* principles, it is also *a priori*. It emphasizes that the words/sentences would have no meaning outside that whole and would have a different meaning in a different whole. The PC mentions syntax and applies to complex expressions. SH applies to both simple and complex, structured expressions as well. It is also very general and is not tied to a specific notion of meaning, view on human language, or program in semantics. It simply requires sentences and truth conditions, as opposed to words and referents, to be basic in semantics. SH is also a metaphysical view on language and meaning (see Fodor and Lepore, 1992). In that respect, it has no clear purchase value in theoretical or applied linguistics. On the other hand, as long as it is not incompatible with empirical research, there is no point in rejecting (or in accepting) it.

One can argue for SH in Conceptual Role Semantics (Block, 1986). That conception of semantics is grounded on the functionalist theory of the mind, which takes thought contents to be functionally defined and the meaning of sentences/words to reflect thought contents (*see* **Mentalese**).

According to Conceptual Role Semantics, the meaning of a token of a linguistic expression is constituted by its inferential properties.

In this view, SH does not follow from methodological principles, but is a plain assertion about meaning. It also concerns tokens, not types, and in that respect, it is an individualistic doctrine. This version of SH is arguably a metaphysical approach to meaning. It is also designed to fit a functionalist view of mental states. Block, the main proponent of Conceptual Role Semantic (CRS), was careful to mention that CRS is not a theory, but a framework for developing semantic theories.

Problems with Semantic Holism

SH raises vexing philosophical issues. Does SH reflect substantial properties of meaning? Does it have an ontological scope? If human languages are holistic, is being holistic an essential property of human languages? Or does SH just echo our way of capturing meaning in a theory, being an artifact produced

by our way of theorizing/knowing the semantics of natural languages? SH has *prima facie* implausible consequences for linguistics.

Learning Problem If SH is true, then one cannot learn the meaning of a lexical item without learning a whole language. It also implies that, to understand the meaning of one single word/sentence a speaker utters, one must be able to understand all words/sentences that speaker can utter. To master the meaning of one single word, one must master a whole language. That makes learning and understanding a language an 'all or nothing' affair, and it implies that learning/mastering part of a language does not make sense. Friends of holism reply that "in so far as we take the 'organic' nature of language seriously, we cannot accurately describe the first steps toward its conquest as learning part of a language; rather, it is a matter of partly learning" (Davidson, 1965: 7). Some (Dressner, 2002) have argued that this fits what we know about the psychology of language learning. Holism apparently opposes the PC – the meaning of one sentence cannot be determined in the way described by the PC because the whole language is involved. However, it is arguable that there is no problem in that quarter. The PC presupposes that the meaning of an unstructured item is determined, and a functional relation is applied to these items. Being determined by and being a function of are two different relationships. Insofar as determination is concerned, *prima facie*, both holism and the PC can be true.

Instability Problem SH also implies that, by adding terms to the language or by altering the meaning of one single word/sentence in a language, one alters the meaning of all words/sentences in that language. As a consequence, the language of a speaker changes constantly because words/sentences are introduced and/or the meaning of lexical items is altered. In addition, two speakers do not speak the same language or use expressions having the same meaning, because the meaning of the words/sentences of a speaker differs from the meaning of the words/sentences of another speaker insofar as they do not share the same vocabulary or produce the same sentences.

Disagreement Problem If two tokens of the same type differ in meaning, then speakers cannot disagree or agree, for that matter, on a sentence. The fact that you accept and that I reject 'Snow is white' does not show that we disagree on anything; the fact that we both accept the same sentence does not show that we agree on anything. The reason is the same in both cases: The tokens have different meaning in the two idiolects, and agreement or disagreement requires sameness of meaning.

SH can be escaped if truth and sentences are not treated as basic in semantic. It can also be avoided if the PC or the CP is rejected. The PC draws wide agreement. The CP is more controversial, and Atomism rejects it. SH inherits certain indeterminacy from the premises leading to it. What do 'is determined by,' 'depends on,' 'is constituted by,' and 'meaning' mean exactly in SH? These terms are not designed to be read in a mathematical way, and they remain controversial. How can sound, *a priori* methodological principles have dramatic metaphysical consequences? Those questions do not make the problem disappear, but prove how opaque it is. The Conceptual Role Semantic version of Semantic Holism can be rejected by refusing CRS.

Semantic Holism and the Philosophy of Mind

Meaning is connected to thought content. Replace the relevant words in the above principles, and you end up with the idea that thought contents have a compositional aspect and that a single thought content is determined by all of one's thought contents. Changing one belief also implies altering the web of beliefs, and hence each belief – this is the instability problem raising its head in the philosophy of mind – and to know one single thought content of a speaker, one must know all of his other thought contents. Fodor and Lepore (1992) drew attention to the fact that, if SH is true and if meaning is echoed in thought contents, then we cannot share a thought unless we share all of our thought contents, and trying to state psychological laws is a futile project, doomed from the start.

Conclusion

Wherever one looks, one finds only implausible consequences of SH. As a consequence, it never served as the basis for a detailed research program in semantics. However, it did bring the analytic philosopher one step closer to the Hermeneutic tradition. Hermeneutics is thoroughly holistic.

See also: Compositionality: Philosophical Aspects; Context Principle; Empiricism; Mentalese; Semantic Value; Thought and Language: Philosophical Aspects; Truth Conditional Semantics and Meaning; Verificationism.

Bibliography

Block N (1986). 'Advertisement for a semantics for psychology.' *Midwest Studies in Philosophy,* 615–678.

Davidson D (1965). 'Theories of meaning and learnable languages.' In Davidson D (ed.) (1980). *Inquiries into truth and interpretation.* Oxford: Clarendon Press.
Davidson D (1967). 'Truth and meaning.' In Davidson D (ed.) (1984). *Inquiries into truth and interpretation.* Oxford: Clarendon Press.
Davidson D (1984). 'Truth and meaning.' In Davidson D (ed.) *Inquiries into truth and interpretation.* Oxford: Clarendon Press.
Dressner E (2002). 'Holism, language acquisition, and algebraic logic.' *Linguistics and Philosophy 25,* 419–452.
Fodor J & Lepore E (1992). *Holism. A shopper's guide.* Oxford: Basil Blackwell.
Frege G (1978). *The foundations of arithmetic.* Oxford: Basil Blackwell.
Hempel C G (1950). 'The empiricist criterion of meaning.' In Ayer A J (ed.) *Logical positivism.* The Free Press. 108–129.
Okasha S (2000). 'Holism about meaning and about evidence: in defence of WV Quine.' *Erkenntnis 52,* 39–61.
Quine W V O (1980). 'Two dogmas of empiricism.' In Quine W V O (ed.) *From a logical point of view.* Cambridge, MA: Harvard University Press.
Quine W V O (1986). 'Reply to Vuillemin.' Hahn L E & Schilpp P A (eds.) *The philosophy of W. V. Quine.* La Salle: Open Court.

Holmes, Janet (b.1947)

M Meyerhoff, University of Edinburgh, Edinburgh, UK

Janet Holmes was born in Liverpool on May 17, 1947. The eldest of Muriel and Robert Quirk's four children, she completed a B.A. at Leeds University and then an M.Phil. under the supervision of John Pride and John Strover. In July 1970 she moved to New Zealand to take up a lecturer (assistant professor) position in linguistics at Victoria University of Wellington. She settled there permanently, raising two sons with her husband, Tony. She played an influential role in shaping the character of the Department of Linguistics (formally instituted in 1985). In 1993, Holmes was awarded a personal chair in Linguistics, and in 1994 was elected Fellow of the Royal Society of New Zealand, an honor only rarely extended to social scientists.

Holmes has made New Zealand a dynamic center for research into sociolinguistics. Her skills at establishing collaborative ventures across universities and between the university and the business community have maintained the vitality of the field even in an era of increased competition for decreased academic funding in New Zealand. Her interests are politeness theory, language and gender, ethnicity, discourse, and New Zealand English. Her track record of gaining research funding is impressive; her research projects have not only produced many scholarly publications, but also provided invaluable training for younger scholars, many of whom have gone on to pursue academic careers of their own. Corpus-based social dialectology led by Holmes has made NZ English one of the best documented varieties in the world.

Holmes was one of the first researchers to look critically at Robin Lakoff's notion of 'women's language,' and her work (1984 forward) has highlighted the interplay between gender and other social factors such as politeness. Her work (e.g., 1986) consistently stresses the importance of going beyond assumptions of one-to-one mappings between form and function and supports this with naturally occurring data. Her focus on form–function relations includes careful attention to intonation, still often neglected. She continues to defend the practice of analyzing politeness using both qualitative and quantitative methods, in the face of postmodern challenges to this approach.

Holmes's studies of workplace interaction have ranged from the interactional function of humor, through to analyses of effective management style in medicine, manufacturing, and white collar offices. Here, too, Holmes has stressed interdisciplinarity and action-based research. She is author or editor of 12 monographs, over 180 academic articles, and more than 30 general audience articles on sociolinguistics. Many have been reprinted and translated. She is also an active reviewer for journals and monographs. She has supervised or examined 80 postgraduate theses and dissertations.

See also: Applied Linguistics in Australasia and the Pacific; Corpora in Studies of Variation; Critical Applied Linguistics; Gender; Gender and Language; Language in the Workplace: Different Approaches; Politeness; Prestige, Overt and Covert; Variation and Language: Overview.

Bibliography

Holmes J (1984a). *Bilingual education.* Wellington: Victoria University of Wellington, English Language Institute.
Holmes J (1984b). 'Hedging your bets and sitting on the fence: some evidence for hedges as support structures.' *Te Reo 27,* 47–62.

Davidson D (1965). 'Theories of meaning and learnable languages.' In Davidson D (ed.) (1980). *Inquiries into truth and interpretation.* Oxford: Clarendon Press.
Davidson D (1967). 'Truth and meaning.' In Davidson D (ed.) (1984). *Inquiries into truth and interpretation.* Oxford: Clarendon Press.
Davidson D (1984). 'Truth and meaning.' In Davidson D (ed.) *Inquiries into truth and interpretation.* Oxford: Clarendon Press.
Dressner E (2002). 'Holism, language acquisition, and algebraic logic.' *Linguistics and Philosophy 25,* 419–452.
Fodor J & Lepore E (1992). *Holism. A shopper's guide.* Oxford: Basil Blackwell.
Frege G (1978). *The foundations of arithmetic.* Oxford: Basil Blackwell.
Hempel C G (1950). 'The empiricist criterion of meaning.' In Ayer A J (ed.) *Logical positivism.* The Free Press. 108–129.
Okasha S (2000). 'Holism about meaning and about evidence: in defence of WV Quine.' *Erkenntnis 52,* 39–61.
Quine W V O (1980). 'Two dogmas of empiricism.' In Quine W V O (ed.) *From a logical point of view.* Cambridge, MA: Harvard University Press.
Quine W V O (1986). 'Reply to Vuillemin.' Hahn L E & Schilpp P A (eds.) *The philosophy of W. V. Quine.* La Salle: Open Court.

Holmes, Janet (b.1947)

M Meyerhoff, University of Edinburgh, Edinburgh, UK

Janet Holmes was born in Liverpool on May 17, 1947. The eldest of Muriel and Robert Quirk's four children, she completed a B.A. at Leeds University and then an M.Phil. under the supervision of John Pride and John Strover. In July 1970 she moved to New Zealand to take up a lecturer (assistant professor) position in linguistics at Victoria University of Wellington. She settled there permanently, raising two sons with her husband, Tony. She played an influential role in shaping the character of the Department of Linguistics (formally instituted in 1985). In 1993, Holmes was awarded a personal chair in Linguistics, and in 1994 was elected Fellow of the Royal Society of New Zealand, an honor only rarely extended to social scientists.

Holmes has made New Zealand a dynamic center for research into sociolinguistics. Her skills at establishing collaborative ventures across universities and between the university and the business community have maintained the vitality of the field even in an era of increased competition for decreased academic funding in New Zealand. Her interests are politeness theory, language and gender, ethnicity, discourse, and New Zealand English. Her track record of gaining research funding is impressive; her research projects have not only produced many scholarly publications, but also provided invaluable training for younger scholars, many of whom have gone on to pursue academic careers of their own. Corpus-based social dialectology led by Holmes has made NZ English one of the best documented varieties in the world.

Holmes was one of the first researchers to look critically at Robin Lakoff's notion of 'women's language,' and her work (1984 forward) has highlighted the interplay between gender and other social factors such as politeness. Her work (e.g., 1986) consistently stresses the importance of going beyond assumptions of one-to-one mappings between form and function and supports this with naturally occurring data. Her focus on form–function relations includes careful attention to intonation, still often neglected. She continues to defend the practice of analyzing politeness using both qualitative and quantitative methods, in the face of postmodern challenges to this approach.

Holmes's studies of workplace interaction have ranged from the interactional function of humor, through to analyses of effective management style in medicine, manufacturing, and white collar offices. Here, too, Holmes has stressed interdisciplinarity and action-based research. She is author or editor of 12 monographs, over 180 academic articles, and more than 30 general audience articles on sociolinguistics. Many have been reprinted and translated. She is also an active reviewer for journals and monographs. She has supervised or examined 80 postgraduate theses and dissertations.

See also: Applied Linguistics in Australasia and the Pacific; Corpora in Studies of Variation; Critical Applied Linguistics; Gender; Gender and Language; Language in the Workplace: Different Approaches; Politeness; Prestige, Overt and Covert; Variation and Language: Overview.

Bibliography

Holmes J (1984a). *Bilingual education.* Wellington: Victoria University of Wellington, English Language Institute.
Holmes J (1984b). 'Hedging your bets and sitting on the fence: some evidence for hedges as support structures.' *Te Reo 27,* 47–62.

Holmes J (1984c). '"Women's language": a functional approach.' *General Linguistics 24(3),* 149–178.
Holmes J (1986). 'Functions of *you know* in women and men's speech.' *Language in Society 15,* 1–21.
Holmes J (1995). *Women, men and politeness.* London and New York: Longman.
Holmes J (1997a). 'Maori and Pakeha English: some New Zealand social dialect data.' *Language in Society 26(1),* 65–101.
Holmes J (1997b). 'Women, language and identity.' *Journal of Sociolinguistics 2(1),* 195–223.
Holmes J (2001). *An introduction to sociolinguistics* (2nd edn.). London: Longman. (1st edn. 1992).
Holmes J & Meyerhoff M (eds.) (2003). *The handbook of language and gender.* Oxford: Blackwell.
Holmes J & Stubbe M (2003). *Power and politeness in the workplace: a sociolinguistic analysis of talk at work.* London and New York: Longman.

Holtsmark, Anne (1896–1974)

S Kürschner, Albert-Ludwigs-Universität, Freiburg im Breisgau, Germany

On June 6, 1896, Anne Elisabeth Holtsmark was born in Oslo, Norway. She started her scientific career with university studies in Norwegian, French, and history in Oslo, where she finished her studies in 1924. From 1925 to 1926 she was lector in Hamburg, Germany. She returned to Oslo in 1926 and worked as an assistant at the University Library. In this position, she started extensive philological work including linguistics, modern literary studies, and scientific research on Old Norse manuscripts from Iceland and Norway. Her work included a considerable number of translations to New Norwegian. Anne Holtsmark became a philologist with international influence in the field of Scandinavian studies.

Holtsmark published her work on the manuscript AM 114a 4° in 1931 with the title 'En tale mot biskopene' (A speech against the bishops). Her text-critical work included new results on the potential author of this important Norwegian manuscript from the Middle Ages, which defended the Norwegian king's steps against the church. By palaeographical and linguistic contrast, Holtsmark shed doubt on earlier ideas concerning the exact date of the emergence and the author. Her methods included new advances in physical manuscript treatment: Holtsmark introduced ultraviolet wave illumination to make greased parts of the manuscript readable. By adding the manuscript in facsimile and in a diplomatory format, Anne Holtsmark made the text available to a greater public. Editing Norse manuscripts continued to be one of her main scientific fields in the future, cf. the facsimile editions of 'A Book of Miracles' (AM 645 4°), *Legendarisk Olavssaga* (The legendary Olavssaga) (DelaGardie 8 II), and the *Olav Tryggvasons saga* (AM 310 4°).

From 1931, Holtsmark was university lecturer of Old Norse philology in Oslo. From that moment on, she started university teaching parallel to her research work. Her teaching concentrated on the literature of the Nordic Middle Ages, most of all the Edda and the sagas. In 1936, Holtsmark received a doctoral degree for her work on the 'First grammatical treatise,' a manuscript from the 12th century on the Icelandic grammar of that time. Her dissertation was entitled 'En islandsk scholasticus fra det 12. århundrede' (An Icelandic scholar from the 12th century). Holtsmark offered new insights into the question of who the author of this first Nordic grammar was. Holtsmark studied the manuscript's origin using new methods: she compared the author's ideas to the Western European scientific traditions of that time, especially concentrating on France and the possible Irish influence (which she showed to be overestimated). Her conclusion on the 'First grammatical treatise' was that the grammar featured an exact description of the Old Icelandic sound system. This was carefully worked out to form an orthophonic alphabet. Holtsmark continued her work on old manuscripts from a literary and historical point of view. In addition, her research topics began to include the myths and the legends in the skald and Edda literature.

In 1949, Anne Holtsmark became professor of Old Norse philology at the University of Oslo. She continued her linguistic work and prepared a collection of the vocabulary in the oldest Norwegian manuscripts. It resulted in the book *Ordforrådet i de eldste norske håndskrifter til ca. 1250* (The vocabulary in the oldest Norwegian manuscripts up to c. 1250), which was published in 1955. The work had been started in the early 20th century by Swedish philologists. It was planned to be a companion volume to a similar edition on the Old Icelandic vocabulary by Ludvig Larsson published in 1891. Holtsmark's own editing of Old Norwegian manuscripts made the entire collection more effective. When the project moved to her institute in Oslo, the collection was finally completed.

Anne Holtsmark wrote articles in large numbers, mostly for Norwegian philological journals like *Maal*

Holmes J (1984c). '"Women's language": a functional approach.' *General Linguistics 24(3),* 149–178.
Holmes J (1986). 'Functions of *you know* in women and men's speech.' *Language in Society 15,* 1–21.
Holmes J (1995). *Women, men and politeness.* London and New York: Longman.
Holmes J (1997a). 'Maori and Pakeha English: some New Zealand social dialect data.' *Language in Society 26(1),* 65–101.
Holmes J (1997b). 'Women, language and identity.' *Journal of Sociolinguistics 2(1),* 195–223.
Holmes J (2001). *An introduction to sociolinguistics* (2nd edn.). London: Longman. (1st edn. 1992).
Holmes J & Meyerhoff M (eds.) (2003). *The handbook of language and gender.* Oxford: Blackwell.
Holmes J & Stubbe M (2003). *Power and politeness in the workplace: a sociolinguistic analysis of talk at work.* London and New York: Longman.

Holtsmark, Anne (1896–1974)

S Kürschner, Albert-Ludwigs-Universität, Freiburg im Breisgau, Germany

On June 6, 1896, Anne Elisabeth Holtsmark was born in Oslo, Norway. She started her scientific career with university studies in Norwegian, French, and history in Oslo, where she finished her studies in 1924. From 1925 to 1926 she was lector in Hamburg, Germany. She returned to Oslo in 1926 and worked as an assistant at the University Library. In this position, she started extensive philological work including linguistics, modern literary studies, and scientific research on Old Norse manuscripts from Iceland and Norway. Her work included a considerable number of translations to New Norwegian. Anne Holtsmark became a philologist with international influence in the field of Scandinavian studies.

Holtsmark published her work on the manuscript AM 114a 4° in 1931 with the title 'En tale mot biskopene' (A speech against the bishops). Her text-critical work included new results on the potential author of this important Norwegian manuscript from the Middle Ages, which defended the Norwegian king's steps against the church. By palaeographical and linguistic contrast, Holtsmark shed doubt on earlier ideas concerning the exact date of the emergence and the author. Her methods included new advances in physical manuscript treatment: Holtsmark introduced ultraviolet wave illumination to make greased parts of the manuscript readable. By adding the manuscript in facsimile and in a diplomatory format, Anne Holtsmark made the text available to a greater public. Editing Norse manuscripts continued to be one of her main scientific fields in the future, cf. the facsimile editions of 'A Book of Miracles' (AM 645 4°), *Legendarisk Olavssaga* (The legendary Olavssaga) (DelaGardie 8 II), and the *Olav Tryggvasons saga* (AM 310 4°).

From 1931, Holtsmark was university lecturer of Old Norse philology in Oslo. From that moment on, she started university teaching parallel to her research work. Her teaching concentrated on the literature of the Nordic Middle Ages, most of all the Edda and the sagas. In 1936, Holtsmark received a doctoral degree for her work on the 'First grammatical treatise,' a manuscript from the 12th century on the Icelandic grammar of that time. Her dissertation was entitled 'En islandsk scholasticus fra det 12. århundrede' (An Icelandic scholar from the 12th century). Holtsmark offered new insights into the question of who the author of this first Nordic grammar was. Holtsmark studied the manuscript's origin using new methods: she compared the author's ideas to the Western European scientific traditions of that time, especially concentrating on France and the possible Irish influence (which she showed to be overestimated). Her conclusion on the 'First grammatical treatise' was that the grammar featured an exact description of the Old Icelandic sound system. This was carefully worked out to form an orthophonic alphabet. Holtsmark continued her work on old manuscripts from a literary and historical point of view. In addition, her research topics began to include the myths and the legends in the skald and Edda literature.

In 1949, Anne Holtsmark became professor of Old Norse philology at the University of Oslo. She continued her linguistic work and prepared a collection of the vocabulary in the oldest Norwegian manuscripts. It resulted in the book *Ordforrådet i de eldste norske håndskrifter til ca. 1250* (The vocabulary in the oldest Norwegian manuscripts up to c. 1250), which was published in 1955. The work had been started in the early 20th century by Swedish philologists. It was planned to be a companion volume to a similar edition on the Old Icelandic vocabulary by Ludvig Larsson published in 1891. Holtsmark's own editing of Old Norwegian manuscripts made the entire collection more effective. When the project moved to her institute in Oslo, the collection was finally completed.

Anne Holtsmark wrote articles in large numbers, mostly for Norwegian philological journals like *Maal*

og Minne. She wrote and edited 10 scientific monographs, including work on texts from the Middle Ages, all within her broad field of research. In addition, Anne Holtsmark wrote 120 articles for the *Kulturhistorisk leksikon for nordisk middelalder* (Culture historical lexicon for the Nordic Middle Ages). She was one of the founders of the Old Norse Dictionary Association (Gammelnorsk Ordboksverk) and the Association for Norwegian Culture Research (Forening til Norsk Kulturgransking). From 1941, Holtsmark was a member of the Norwegian Scientific Academy. She received the Royal Norwegian Order of Saint Olav and the Icelandic Falke-order for her lifetime work.

Holtsmark's main interest was the medieval language and its sources in literary texts, and she ranks among the most important Scandinavian philologists, especially for medieval studies. In her scientific work, Holtsmark maintained a broad scope, including linguistics, literature studies, and broader cultural studies. She also worked in fields combining all disciplines, and she presented results not possible without considerable knowledge of all these fields. In all of her work, Holtsmark was known to be very accurate. She based her results on a very exact text analysis, be it literary or linguistic research. For this reason, Holtsmark's work is still of great importance for current medieval research in Scandinavian Studies. Anne Holtsmark died on May 19, 1974.

See also: Norse and Icelandic; Norwegian.

Bibliography

Hødnebø F (1975). *Anne Elisabeth Holtsmark. 21/6/1896–19/5/1974. Arkiv för Nordisk Filologi 90,* 203–234.

Holtsmark A (1931). *En tale mot biskopene. En sproglighistorisk undersøkelse.* Skrifter utgitt av Det Norske Videnskaps-Akademi i Oslo, II. hist.-filos. klasse 1930, no. 9. Oslo: Dybwad.

Holtsmark A (1936). *En islandsk scholasticus fra det 12. århundrede.* Skrifter utgitt av Det Norske Videnskaps-Akademi i Oslo, II. Hist.-Filos. Klasse 1936, no. 3. Oslo: Dybwad.

Holtsmark A (ed.) (1955). *Ordforrådet i de eldste norske håndskrifter til ca. 1250* (The vocabulary in the oldest Norwegian manuscripts up to c. 1250). Utgitt av Gammelnorsk Ordboksverk. Oslo: Dybwad.

Homans, George C. (1910–1989)

W A Hass, St Peter, MN, USA

George Homans was a pioneer of sociological exchange theory. Born in Boston, in the Back Bay area, Homans was the son of a prominent physician; his mother was a member of the famous Adams clan. He attended the St Paul's School, Concord, New Hampshire, and Harvard (B.A., 1932). He became well acquainted with Pareto's work just after graduation, and took up sociological investigations as a Harvard junior fellow. His teaching career began and continued in sociology at Harvard (1939–1980), with time out during World War II to serve as a naval officer, returning to the Department of Social Relations there in 1945 until the end of his scholarly career.

Homans devoted his intellectual efforts to explaining social behavior, or, in slightly different terms, to basing social sciences on propositions from behavioral psychology. Language plays no particular overt role in his conceptualization; rather than focus on verbal communication he turned to interpersonal interaction, especially exchange. His fundamental principle is that of instrumental conditioning, or less psychologically, the social feedback loop. Accordingly, in his 1961 *Social behavior*, his first general proposition is: "If in the past the occurrence of a particular stimulus-situation has been the occasion on which a man's activity has been rewarded, then the more similar the present stimulus-situation is to the past one, the more likely he is to emit the activity, or some similar activity, now" (p. 53).

His analysis of human interaction, which can be seen as a modern version of hedonism or utilitarianism, is found implicitly in his *English villages of the 13th century* (1941). It is demonstrated explicitly, but still predominantly inductively in his *Human group* (1950), which approaches various specific face-to-face social phenomena in behavioral terms. These include the Western Electric Hawthorne Works in Chicago, an East Coast slum gang, Tikopia families, a New England village, and an electrical equipment company. These groupings allow him to point out issues in social organization, which offer paradigms for understanding society on a larger, institutional and cultural scale as well.

Homans then went on to formulate more general behavioral principles, in terms of rewarded activities, and the exchanges in which they are involved, and to

L

Laboratory Phonetics

J Kingston, University of Massachusetts, Amherst, MA, USA

Anyone who has seen *My Fair Lady* will recall the scene in which Henry Higgins sets Eliza Doolittle the task of pronouncing "In Hartford, Hereford, and Hampshire, hurricanes hardly ever happen" to test her dropping of [h]s where they belong and her adding them where they do not belong. Eliza is to speak this utterance into a cone attached to a tube whose other end is attached to a valve emitting a flame. Every time she pronounces an [h], the flame briefly jets out further – this only occurs when she says "[h]ever" and not when she says any of the words spelled with "h." The other part of this apparatus is a revolving cube with mirrored sides, in which the flame is reflected. Eliza becomes mesmerized by the revolving reflections and mechanically recites the sentence over and over, until the flame ignites her script.

The device depicted is apparently a 'Koenig flame manometer' (Koenig, 1872), which was widely used by phoneticians in the late 19th and early 20th centuries to measure variation in acoustic pressure from the mouth: the flame jets out further when acoustic pressure is higher. The revolving mirrored cube thus captures changes in acoustic pressure over time in the height of the reflected flame. Because the observer does not separate the successive reflections as each face of the cube turns into the plane of view, they are superimposed onto one another, creating an apparently static image of time-varying acoustic pressure. Different speech sounds produce different images, which could then be used to investigate their phonetic properties. (Ironically, this device could not be used to measure much slower changes in aerodynamic pressure, nor the increase in oral air flow that characterizes an [h].)

Another memorable scene from *My Fair Lady* is the one when Eliza first meets Henry Higgins, and he demonstrates his ability to place people in London and elsewhere in England by the nuances of their pronunciation. He also shows Eliza and the others what he's been writing down in his notebook, a phonetic transcription of her speech.

These two scenes illustrate the contrasts between 'field phonetics' and 'laboratory phonetics,' the topic of this entry. The goal of field phonetics is a **qualitative** description, which uses categorically distinct symbols to record the transcribable phonetic properties of a person's speech. To confirm the transcription, a skilled field phonetician typically imitates the speaker's utterance. If the speaker judges the imitation to be a close match, the phonetician infers that both the original utterance and the imitation were pronounced more or less identically. Both the goal and methods of laboratory phonetics are quite different. Its goal is to **quantify** the continuously varying values of the speaker's articulations, the resulting acoustics, and the listener's perceptual response, and to relate those continuous quantities to the categorical qualities that distinguish one message from another. To achieve this goal, laboratory phonetics uses instruments and experiments, rather than the careful listening and imitation of field phonetics. Yet despite the difference in principle between laboratory phonetics and field phonetics, in practice, most phonetic research combines the two, as Henry Higgins did. Indeed, it is more than fitting that the phonetic consultant to *My Fair Lady* was Peter Ladefoged, the phonetician whose work epitomized the combination of field and laboratory phonetics.

The remainder of this entry is broken into three parts. The first part discusses the instrumentation and software that are widely used in laboratory phonetics and describes the uses to which they are put. The second part lists persistent research problems and the research practices that have been developed in attempts to solve them; this list is incomplete. Finally, the third section discusses two recent important developments in laboratory phonetics: the increasing use of natural corpora, and the attempt to integrate phonetics and phonology, called 'laboratory phonology.'

Instrumentation and Software

The Koenig flame manometer was simply one of a large variety of ingenious instruments and techniques devised in the late 19th and early 20th centuries to measure phonetic properties of utterances. Among many others, the reports of Rousselot (1897–1908) and Scripture (1902) stand out as comprehensive reviews of their own and many others' instrumental investigations of virtually all aspects of the articulation, acoustics, and perception of speech sounds. In many cases, today's instruments and software are merely refinements of devices invented a century and more ago.

Aerodynamics

Speech is the transformation of energy first embodied in air flow into sound, so this review of instrumentation begins with the devices used to measure changes in the volume and pressure of the cavities containing air and air flow between these cavities, i.e., the aerodynamics of speech.

Changes in chest cavity volume are measured with either a 'plethysmograph' or transducers attached to the chest wall and abdomen. In a whole-body plethysmograph, the speaker is placed in a large airtight container and breathes in and out through a mask connected to the outside. As the speaker's chest and abdomen expand and contract, his or her body takes up more and then less room in the container, compressing and then rarefying the surrounding air. These changes in pressure directly reflect the changes in chest cavity volume. In a variant design, compressible air bulbs are attached to the speaker's chest and abdomen, and chest cavity volume is measured by the changes in air pressure inside these bulbs. An alternative means of measuring changes in chest cavity volume is to attach a magnetic field source to the back and transducers capable of measuring the strength of that field to the front of the chest and abdomen; as the chest and abdomen expand and contract, the distance between the field source and the transducer increases and decreases, and with those distance changes, the strength of the field detected by the transducer wanes and waxes.

These volume changes are all intended to change air pressure in the lungs, which in turn determines whether air flows into or out of the lungs. Speech is nearly always produced during an exhalation, when the pressure in the lungs is greater than the pressure downstream in the vocal tract. The pressure of the air in the lungs is also a measure of how much energy is available to produce sound. This pressure is usually measured in the trachea, either through a pressure transducer attached either to a wide-bore needle inserted between two tracheal rings, or via a catheter attached to a balloon swallowed into the esophagus (the esophagus shares a wall with the trachea, so as pressure rises in the trachea, distending its walls, the balloon in the esophagus is compressed). The energy embodied in the tracheal pressure is first transformed into sound when it causes the vocal folds to vibrate, producing a rapid variation in acoustic pressure just downstream. Speakers contract the lungs more rapidly when emphasizing a syllable, and the resulting increase in air pressure in the trachea increases both the rate and size of vocal fold vibration and thus the fundamental frequency and intensity of this sound source (Ohala, 1990).

Air pressure in the oral cavity is measured via a catheter inserted either into the corner of the mouth or through a nostril over the soft palate into the pharynx. The pressure transducer may be in the catheter itself or outside. Some sounds obstruct the flow of air out of the mouth more than others and cause air pressure to build up behind that obstruction. When air pressure is elevated, air flow increases through any open channel. If the increase is great enough, flow becomes turbulent, and this turbulence is audible as noise. This is the second way in which energy embodied in elevated pressure is turned into sound.

Finally, air flow out of the mouth or nostrils, or even inside the oral cavity, is measured using a 'pneumotachograph.' This device applies Poiseuille's law, which states that laminar (nonturbulent) flow between two locations in a tube is proportional to the pressure loss between those locations. In practice, a known resistance to airflow is introduced between the two locations at which pressure is measured.

Muscle Activity

The volume of the chest cavity could not be changed, nor could any other articulation be made, without contracting muscles. Because voltages change in muscles when they contract, the extent and timing of their contraction can be measured by attaching electrodes to the skin surface overlying the muscles or actually inserting the electrodes directly into the muscles. Surface electrodes capture the voltage changes in all the underlying muscle fibers, as do large electrodes inserted into the muscles themselves, but it is also possible to record from single muscle units using finer electrodes. This technique is called 'electromyography.'

Glottal Articulations

The larynx, or more specifically, the vocal folds serve as a valve controlling air flow between the lungs or

subglottal cavity and the oral or supraglottal cavity (the glottis is the space between the vocal folds through which air flows when this valve is open). As already noted, the vibration of the vocal folds is also the one of the two means whereby the energy in the air flowing out of the lungs is transformed into sound.

The position and motion of the vocal folds can be imaged directly by inserting a mirror through the mouth into the upper pharynx, or via an endoscope inserted into the mouth or through the nose into the pharynx. The motion of the vocal folds during vibration can then be captured by filming the image at high speeds. When the light source is strobed synchronously with vocal fold vibration, it is also possible to capture still images of the folds at different moments in their vibratory cycle.

'Electroglottography' (EGG) and 'photoglottography' (PGG) are two other means of measuring the position of the folds and the state of the glottis. In EGG, two electrodes are placed on either side of the thyroid cartilage at the height of the glottis and a weak current is passed between them. As the folds move apart or together, opening or closing the glottis, resistance to this current's flow increases or decreases. In PGG, a light source is placed above or below the glottis, and a phototransducer is placed on the other side. As the folds move apart or together, the amount of light passing through the glottis and falling on the transducer waxes and wanes. These two techniques complement one another. The change in resistance to current flow is greatest at the moment when the folds first make contact or separate all along their length, so EGG measures the moment when the glottis closes or starts to open most accurately. The most light falls on the phototransducer when the folds separate maximally, so PGG measures the moment of peak opening in a particular glottal articulation with the greatest precision.

All these methods are limited in what they show about the motion or position of the vocal folds: filming only shows the folds from above, and EGG and PGG only depict accurately the moments of fold contact and peak glottal opening. By permitting the folds' cross-section to be imaged, ultrasound laryngography has shown how vibration propagates from the lower to the upper edge of the folds. These limitations of all these techniques apply only to the vocal folds *in vivo*. Larynges excised from dogs have been used extensively to study the nature of vocal fold movement during vibration (Moore and Berke, 1988; Perlman and Titze, 1988; Titze *et al.*, 1997).

Finally, a number of devices have been used to measure the vertical movement of the larynx in the neck. Among them is the 'thyroumbrometer,' in which a light is shone past the profile of the neck, casting a shadow onto a two-dimensional grid of phototransducers on the other side. By recording changes over time in the amount of light falling onto the transducers, the shadow can be tracked and the height of the larynx determined from the position of the thyroid prominence.

Oral Movements and Articulatory Contact

Movements of the oral articulators – the tongue, lips, and soft palate – have been measured mechanically and by various imaging techniques. The mechanical means all involve attaching to or resting a low-mass lever on the articulator of interest and then measuring the movement of its other end when that articulator moves. In the past, the movement of the other end was measured directly using a kymograph, a device that draws a trace on paper that is proportional in size to the lever's movement. In modern use of this technique, the other end of the lever is instead attached to a strain gauge, which emits a voltage proportional to the size of the movement of the lever arm. This method is most suitable for measuring movements of the visible articulators – the lips and jaw – and has also been used successfully in recording soft palate movements (the lever is inserted through the nose and one end rests on top of the soft palate), but it is not suitable for measuring tongue movement, because this articulator's shape and position in the oral cavity change enormously from one speech sound to the next. An imaging technique has also been widely used to track the movements of the visible articulators: Infrared-emitting diodes are placed on the articulators of interest, and their positions in three dimensions are tracked by combining images from multiple cameras (Munhall and Vatikiotis-Bateson, 1998).

Filming of radiographs ('cineradiography') was the first means used to image tongue movement effectively. The tongue and any other oral articulator of interest would be painted with a barium paste to improve contrast, and the X-ray beam would be directed at the speaker from the side. The disadvantage of this technique is that the speaker is exposed to an enormous amount of ionizing radiation in a very short period of time, so only a small amount of data can be collected. Because even the radiation dose received in a short recording session is dangerous to the speaker, this technique is no longer used.

The 'X-ray microbeam' was built to replace cineradiography (Nadler *et al.,* 1987). This technique aims a very narrow beam of very high-energy X-rays at gold pellets glued to the articulators; the narrowness and high energy of the beam reduces the radiation dose received during a 10-minute recording session (48 mRad) to little more than that received during a

routine full-mouth dental X-ray examination (10 to 30 mRad). Unlike cineradiography, which provides an image of the entire surface of the tongue or any other articulator within the X-ray beam, the X-ray microbeam measures only the positions of the gold pellets, or the 'flesh points' to which they are attached. (The use of cameras to track the positions of infrared-emitting diodes also measures the positions of flesh points.) Fortunately, the tongue is hydrostatic, so the distances between pellets and the corresponding flesh points are not changed by movements of the tongue or changes in its shape.

The other widely used technique for measuring flesh points inside the vocal tract, 'electromagnetic midsagittal articulometry' (EMMA), uses small magnets instead of gold pellets (Perkell *et al.,* 1992; for comparison of the X-ray microbeam and EMMA, see Byrd *et al.,* 1999). The three-dimensional positions of these magnets is registered by an array of magnetic field transducers mounted on a cage surrounding the speaker's head. As implied by its name, EMMA only permits imaging of flesh point locations in the midsagittal (profile) plane. This is also true of the traditional X-ray and X-ray microbeam techniques, because it is too dangerous to direct the X-rays toward either the brain or the spinal cord.

The two remaining techniques, 'ultrasound' and 'magnetic resonance imaging' (MRI), do not share this limitation; they can be used to image the vocal tract in any plane. Nor are these techniques limited to tracking the locations of flesh points; like cineradiography, ultrasound and MRI both produce anatomically complete images. Each, however, is limited in other ways.

Ultrasound images are constructed from the time delay between the sound's emission and the return of its echo from a boundary, e.g., between the flesh and air at the tongue surface (Stone *et al.,* 1988, 1992; Stone and Lundberg, 1996). The delay increases, the further the tongue surface is from the sound source. Although the surface of the tongue root and body can be readily imaged, the surface of the tongue tip and blade often cannot be because there is an air pocket beneath this part of the tongue, and the sound echoes off the boundary between the floor of the mouth and this air pocket before ever reaching the tongue tip and blade.

The principal limitation with MRI is that it takes a long time to record an image. In the earliest uses of this method, the speaker had to move the articulators to their positions in the sound of interest and then hold them there for a number of seconds. Obviously, the continuous movements of connected speech could not be studied under these conditions. However, procedures have recently been developed that permit five to 10 and soon perhaps as many as 20 images to be collected per second (Demolin *et al.,* 2000; Narayanan *et al.,* 2004). In addition to these limitations, visual signal processing algorithms must be applied to these images to quantify articulator positions and movements, a challenge that does not arise in techniques that measure the positions of flesh points.

All of these techniques measure articulatory movements or positions. 'Palatography' measures the contact of the tongue with the palate. Traditionally, a dark substance, e.g., a mixture of charcoal dust and olive oil, is painted on the tongue surface, the speaker pronounces the target sound (nearly always a consonant), a mirror is inserted into the mouth, and the reflection of the palate is photographed. The dark substance rubs off on the palate or teeth in the locations wherever the tongue touched it. The result is a 'palatogram.' A 'linguogram' can be obtained by painting the palate instead and photographing the tongue surface after the speaker pronounces the target sound (see Dart, 1991, 1998 for a recent use of this traditional technique). This technique has two obvious drawbacks: only a single articulation can be measured at a time, and it is impossible to measure its trajectory.

An elaboration of this technique, 'electropalatography' (Gibbon and Nicolaidis, 1999), does away with both of these drawbacks. An artificial palate covered with a dense array of electrodes is fit to the speaker's palate. A circuit is completed whenever the speaker's tongue contacts an electrode, so the pattern and timing of electrode contacts can be used to track the tongue's contact with the palate from moment to moment in indefinitely long stretches of speech. Electropalatography is subject to two limitations: the palates must be fit individually to each speaker, so it is difficult to study more than a few speakers, and the palates do not extend back beyond the hard palate, so it is impossible to record contact made with the tongue body on the soft palate or further back.

Acoustics

The acoustic analysis of speech relies on the source-filter model of speech production (Fant, 1960). A sound source is created when the energy embodied in the flow of air through the vocal tract is transformed into sound. The first kind of sound source is the (quasi-)periodic variation in air pressure produced just above the glottis by the vibrating vocal folds in voiced sounds; trills also produce a (quasi-)periodic sound source just downstream from the vibrating soft palate, tongue tip, or lips. The second kind is the aperiodic pressure variation produced when air rushes

turbulently through a narrow channel in obstruent sounds.

The air inside the vocal tract responds to the energy in these sound sources by resonating at a number of different frequencies, which depend on the articulators' positions. Putting the articulators in one position amplifies or passes energy at some frequencies while attenuating or stopping it at others, and moving the articulators to another position changes the frequencies at which energy is passed or stopped. Because the air resonates at different frequencies depending on the positions of the articulators, positioning the articulators filters the energy in the sound source.

Before the development of digital signal-processing techniques, the principal tools for analyzing the acoustics of speech were the 'oscillograph' and the 'sound spectrograph,' which produce an 'oscillogram' and 'spectrogram,' respectively. The oscillograph is essentially a kind of kymograph, but one that measures the moment-by-moment variation in sound pressure rather than the movement of an articulator. Figure 1 is an oscillogram of the utterance 'laboratory phonetics.' Positive and negative acoustic pressures are displayed about a zero line on the vertical axis in this display, and time is displayed on the horizontal axis. This display is often referred to nowadays as a 'waveform.'

The patterns of pressure variation differ between speech sounds of different manners of articulation and phonation type. During vowels, pressure varies between larger positive and negative extremes than during most consonants; moreover, the pressure variation is periodic in vowels, as expected, given that they are produced with the vocal folds vibrating. Periodic pressure variation can also be seen during the voiced consonants [l], [b], and the two [ɹ]s in 'laboratory,' and during the [n] in 'phonetics.' In place of this periodic pressure variation, the waveform shows brief intervals of silence for the voiceless stops [t] and [k] in 'labora[t^h]ory' and 'phoneti[k]s,' respectively, and brief intervals of aperiodic pressure variation, i.e., noise, for the voiceless fricatives [f] and [s] in '[f]onetik[s].' The interval of noise following the silent interval of the [t^h] in 'labora[t^h]ory' is the aspiration indicated by the superscript '^{h}', which is characteristic of the pronunciation of voiceless stops before stressed vowels in American English. This waveform also displays another regular allophonic variant in this variety of English: the pronunciation of the /t/ in 'phone[ɾ]ics' as a tap. The tap's articulation, a very brief contact of the tongue tip to the alveolar ridge, produces the very brief attenuation of energy observable between the flanking vowels. The waveform shows that vocal fold vibration does not even falter during this consonant's pronunciation, unlike during the /t/ in 'labora[t^h]ory,' where the vocal folds stop vibrating for many milliseconds in the stop closure. Finally, pressure varies much more during the stressed first and third vowels in 'laboratory' and during the stressed second vowel in 'phonetics' than in the unstressed vowels in these words. The stressed vowels are also longer than the unstressed ones.

These observations show that waveforms can be segmented into acoustic intervals. These acoustic intervals do not correspond exactly to the transcribed segments, but nonetheless nearly all the intervals can be assigned to one of these segments – the most difficult segmentations in this example are separating the two [ɹ]s in 'laboratory' from the adjacent vowels. The size of the pressure variation distinguishes vowels

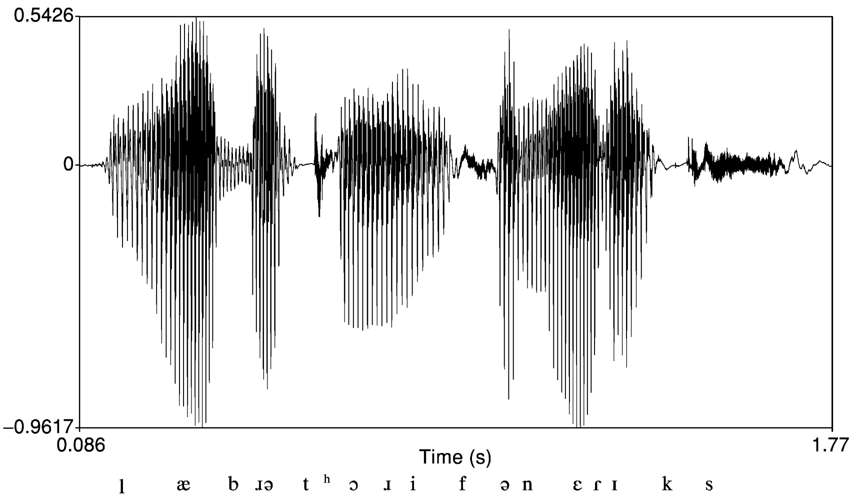


Figure 1 Waveform of the utterance 'laboratory phonetics.'

from consonants, particularly obstruents. The periodic or aperiodic nature of the pressure variation within these intervals can be used to identify their sound sources. Duration can also be measured in both vowel and consonant intervals. However, the durations of these acoustic intervals do not correspond to the durations of the articulatory movements, which begin well before the corresponding acoustic interval and end long after it. So long as the movement is occurring while the vocal tract is not completely closed or narrowly constricted, i.e., not during a stop closure or fricative constriction, then it is possible to see the acoustic effects of these movements by closely inspecting the waveform, as shown by the waveform fragment corresponding to the last three segments [ɔɹi] in 'laboratory' (Figure 2). At the points marked by the arrows, the shapes of the periods change; these changes reflect the changes in resonance frequencies brought about by moving the articulators from their positions in [ɔ] to their positions in [ɹ] and then from [ɹ] to [i].

These changes are much easier to see in a spectrogram of this utterance than the waveform (Figures 3 and 4). The spectrogram displays frequency rather than pressure on the vertical axis, against time on the horizontal axis. The darkness of the mark at a particular frequency and time in the spectrogram indicates how intense the sound is there, i.e., how strongly the air inside the vocal tract is resonating at that frequency and time. The utterance can be divided into acoustic intervals in this display, too, which correspond closely to the intervals discernable in the

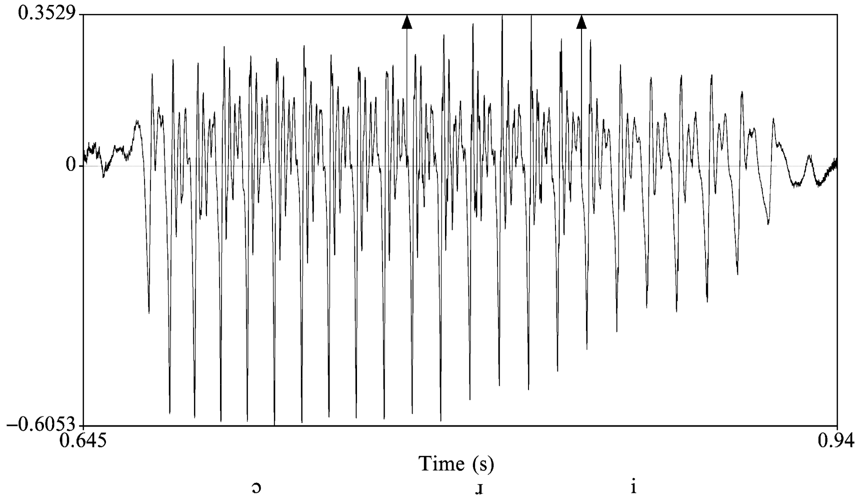


Figure 2 Waveform of the fragment [ɔɹi] from the phrase 'laboratory phonetics.' The arrows mark acoustic discontinuities that roughly correspond to the boundaries between [ɔ] and [ɹ] and between [ɹ] and [i].

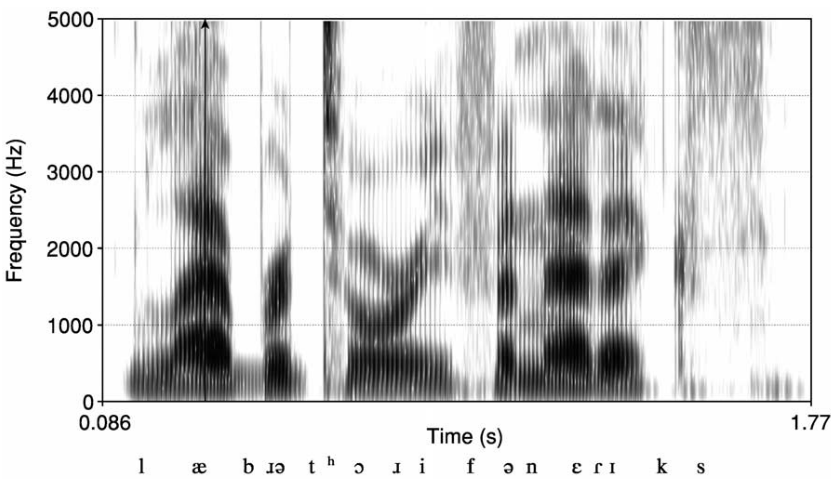


Figure 3 Spectrogram of the phrase 'laboratory phonetics,' where the analysis window is 5 ms wide. The arrow marks the center of the window for the spectrum in **Figure 5**.

waveform in Figure 1. The spectrogram shows that vowels have considerable energy spanning the range from low to high frequencies, while many consonants have much less energy; in some, e.g., the [b] in 'laboratory,' that energy does not extend above the very lowest frequencies. No energy appears, of course, at any frequency during the silent intervals corresponding to the voiceless stop closures of the [tʰ] in 'labora[tʰ]ory' and the [k] in 'phoneti[k]s.' Regular vertical striations can be seen in the intervals during which the speaker is producing a periodic sound source; each one corresponds to a single vibration of the vocal folds. Finally, time-frequency regions in which light and dark areas vary irregularly correspond to intervals of noise, when sound pressure variation is irregular, most obviously in the initial [f] and final [s] in '[f]onetik[s],' as well as the aspiration noise [ʰ] in 'laborat[ʰ]ory.'

The key piece of additional information that is provided by a spectrogram but not a waveform can be seen in the quasihorizontal dark bars visible in all the vowels and some of the consonants. These bars are called 'formants,' and they represent the ranges of frequencies within which the air in the vocal tract resonates strongly when the articulators are in a particular position. During some stretches, especially in the middle of the stressed vowels, these formants are more or less stationary, but at the edges of these stationary intervals, the formants change frequency, often quite rapidly. The stationary intervals, called 'steady-states,' are produced when the articulators are not moving much, while the intervals of formant frequency change, called 'transitions,' are produced when the articulators move from one articulatory target to another. A look at the spectrogram of the stretch corresponding to [ɔɹi]

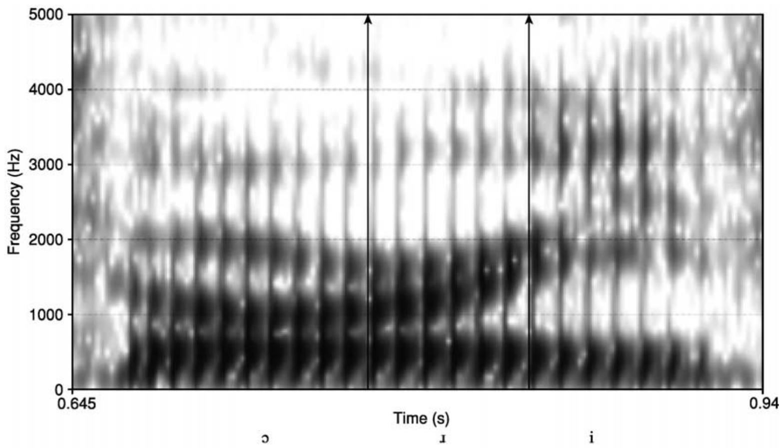


Figure 4 Waveform of the fragment [ɔɹi] from the phrase 'laboratory phonetics.' The arrows mark acoustic discontinuities that roughly correspond to the boundaries between [ɔ] and [ɹ] and between [ɹ] and [i].

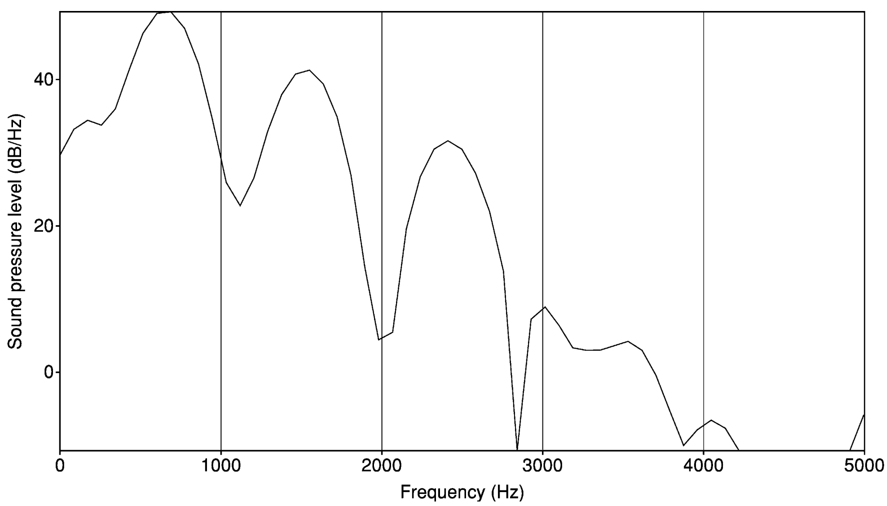


Figure 5 Spectrum of a 5-ms window centered at the arrow in **Figure 3**.

in Figure 4 shows that the formant transitions track the articulatory movements from [ɔ] to [ɹ] to [i] much more noticeably than the subtler changes in the shape of individual periods in the waveform in Figure 2.

The spectrograms in Figures 3 and 4 were produced by computing a fast-Fourier transform (FFT) of successive 5-ms windows of the utterance. A single FFT for just one 5-ms window is displayed in Figure 5. This display is called a 'spectrum,' and it plots amplitude in decibels (dB) on the vertical axis against frequency along the horizontal axis. The time when this spectrum was computed is marked in Figure 3 by an arrow; it is in the middle of the vowel [æ] in the first syllable of 'laboratory.' A comparison with the spectrogram at this moment shows that the peaks in the spectrum correspond to the formants, and the valleys correspond to the frequency ranges in which no formant can be seen.

A spectrogram in which the FFTs were calculated for successive 33.3-ms windows is shown in Figure 6. What is immediately apparent is that, compared to the original spectrogram, the boundaries between successive acoustic intervals are blurred (the vertical striations corresponding to individual cycles of vocal fold vibration have also disappeared), while at the same time the amount of energy present at each frequency is registered with much greater precision. The rule is simple: the finer the temporal resolution of the spectral analysis, the coarser the frequency resolution and vice versa.

A spectrum of the same moment in the middle of the [æ] using this longer window is displayed in Figure 7; it, too, shows the differences in energy levels at each

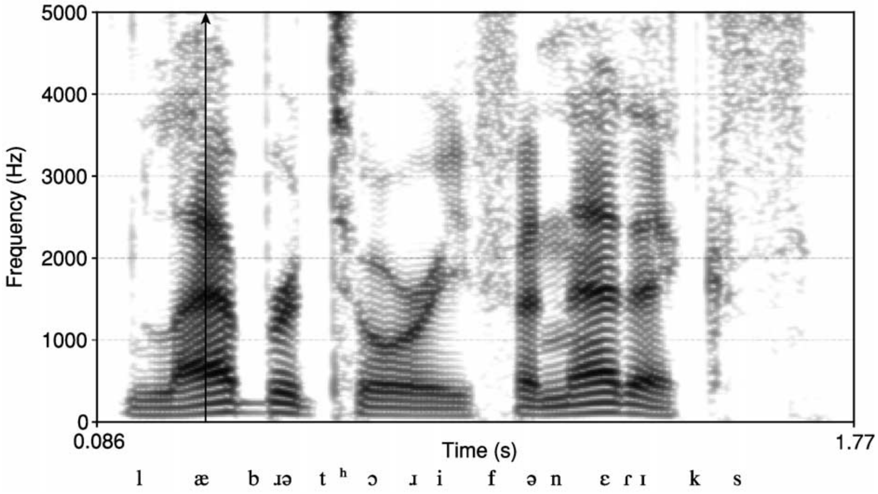


Figure 6 Spectrogram of the phrase 'laboratory phonetics' where the analysis window is 33.3 ms wide. The arrow marks the center of the window for the spectrum in **Figure 7**.

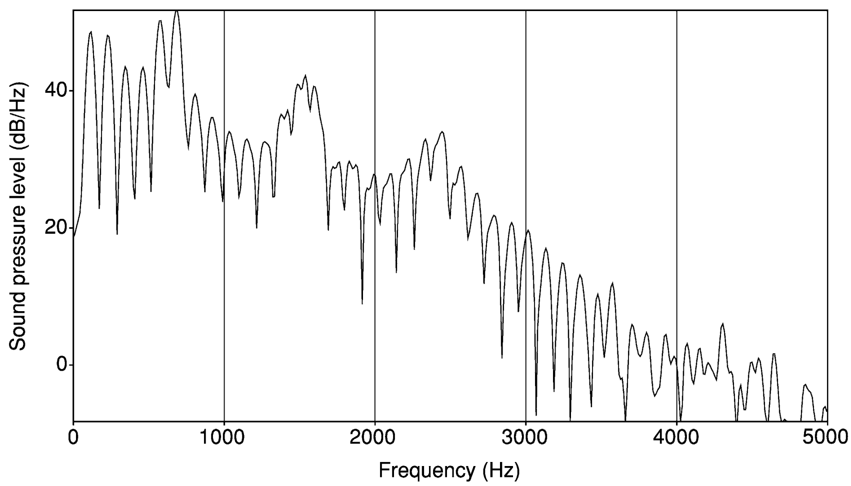


Figure 7 Spectrum of a 33.3-ms window centered at the arrow in **Figure 6**.

frequency much more precisely than did the spectrum in Figure 5. Indeed, the frequency resolution is now fine enough that the individual harmonics of the voice source are separated from one another. In the new spectrogram, the frequencies of the harmonics can be seen to rise and fall with the prosody of the utterance, peaking on the primary stressed syllables and falling to lower values elsewhere (the phrase was spoken with more or less equal prominence on both words). Even so, the coarser image of the energy distribution across frequencies can still be seen easily in the new spectrogram and spectrum. These displays make plain, then, the distinction between source and filter: the frequencies of the harmonics reflect the rise and fall in the voice source's fundamental frequency across the utterance, while their amplitudes reflect the strength of the resonant response of the air in the oral and nasal cavities at each moment in the utterance. The new spectrogram and spectrum show that a formant is produced when the resonant response of the air in the vocal tract amplifies a number of adjacent harmonics.

The spectrograms in Figures 3 and 6 correspond to what are traditionally called 'wide-' and 'narrow-band' spectrograms, so called because the wide-band spectrograms were traditionally made by measuring energy levels in successive 300-Hz-wide windows across the range of frequencies, while narrow-band spectrograms were made using a frequency window that was just 45 Hz wide. The ratio of these frequencies (6.667) is the inverse of the ratio of durations of the time windows (0.15) used in computing the spectrograms and spectra in the figures; thus, the finer the frequency resolution, the coarser the temporal resolution. Spectrograms were traditionally made using a 'spectrograph,' a device that moved filters of one or the other of these widths continuously from the bottom to the top of the frequency range of interest. The energy level within the filters at each successive frequency was transformed into a voltage, whose level determined how dark a mark was burned onto specially treated paper.

Because the harmonics of the voice's fundamental frequency (F0) are resolved in the narrow-band spectrogram, it can be used to measure how fast the vocal folds are vibrating at any moment in the utterance. But this is only one of a myriad of ways in which F0 can be measured; Hess (1983) provided a comprehensive review and evaluation of pitch extraction algorithms. The result of applying one of them – the very widely used auto-correlation method – to the utterance 'laboratory phonetics' is displayed in Figure 8. The other acoustic measure of interest is 'amplitude,' which is calculated by integrating the pressure or intensity values within successive time windows. Its time-varying values for this utterance are displayed in Figure 9.

Formant frequency values could be estimated by eye from the spectrogram in Figure 3; the wide-band spectrogram is most suitable for this purpose because it does not separate the harmonics that are amplified by the air's resonant response. However, it is often desirable to calculate these values automatically. The most widely used method for calculating them is 'linear predictive coding' (LPC) (Markel and Gray, 1976). This method predicts the next value of the signal as the weighted linear combination of some number of preceding values. The weights are chosen so that the difference between the predicted and observed signal value is as small as possible. Their values are the coefficients of a filter that would produce the observed spectrum at that moment in time when applied to a source of known characteristics – a pulse

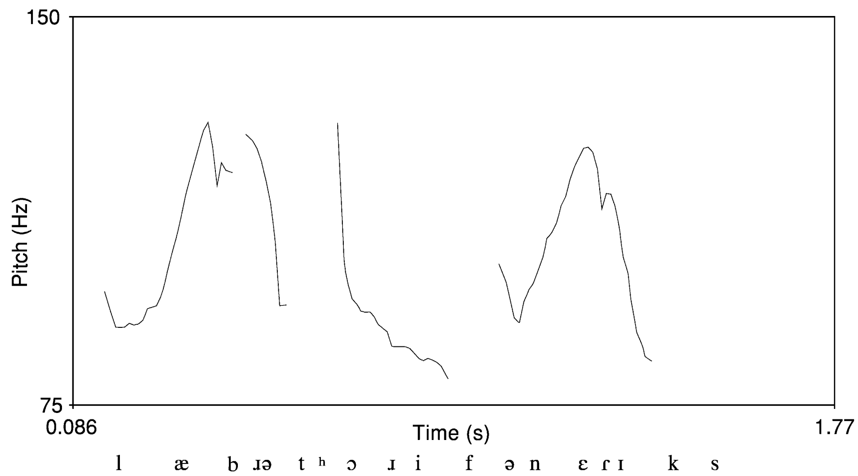


Figure 8 F0 contour of the phrase 'laboratory phonetics.'

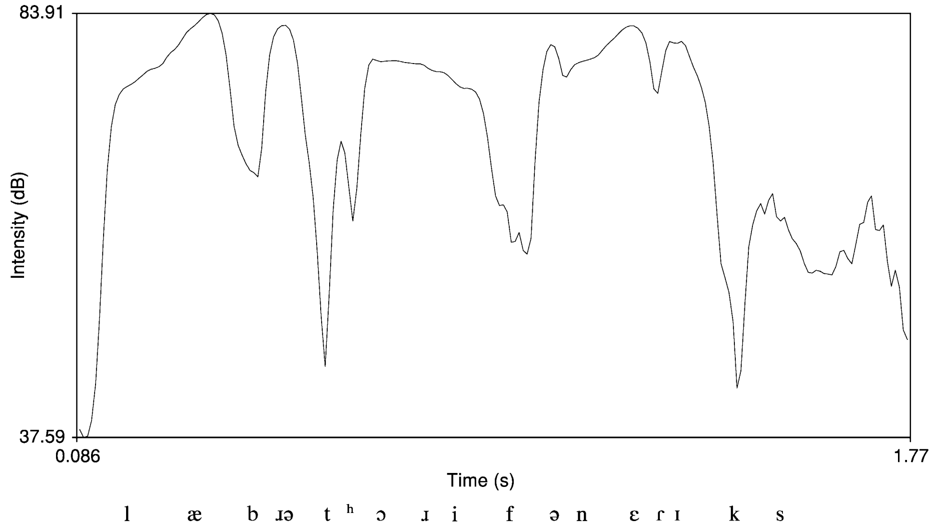


Figure 9 Amplitude contour of the phrase 'laboratory phonetics.'

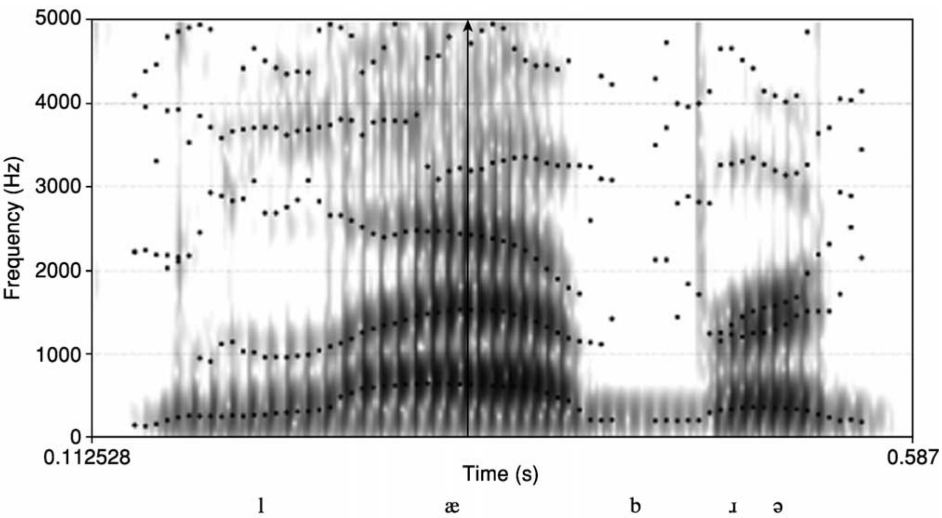


Figure 10 Spectrogram of [læɹbə] from the phrase 'laboratory phonetics' with formants tracked by linear predictive coding analysis superimposed. The arrow marks the center of the window from which the spectra in **Figures 5** and **7** were obtained.

train for periodic (voiced) stretches of the signal and broad-band noise for aperiodic (voiceless) stretches. In other words, the filter models the effects of the vocal tract's resonant response to the energy in the sound source. Because the filter is assumed to specify only the frequencies at which the resonant response amplifies energy in the source, i.e., poles, it estimates the frequencies of the peaks in the spectrum that are candidates to be formants. The number of coefficients chosen determines the number of peaks, and thus the number of formants estimated; their number should generally be greater than twice the number of peaks one wants to find. Because the next signal value is predicted from previous values, this method captures the relatively continuous change in formant frequency values. Estimates of formant frequencies using this method have been superimposed on the formants in the spectrogram of the stretch [læbɹə] from 'laboratory' in Figure 10.

Because LPC predicts the next signal value from the previous signal values, and it does so compactly by decomposing the signal into a filter and source, it is an efficient means of compressing the speech signal for transmission across a channels where a low bit-rate is desirable. For example, it is used in the Global System for Mobile Communication (GSM) standard for voice compression in mobile telephony. After the decomposed signal is transmitted, the original signal

is resynthesized by applying the filter to the source. The resynthesized rendition typically matches the original very closely. Because the signal is decomposed into filter and source characteristics, LPC resynthesis is often used to study the perceptual effects of changing the properties of one of these components while leaving the other unchanged. For example, the F0 values of the source can be changed while leaving the formant frequencies of the filter unchanged to study the perception of intonation in utterances composed of the same consonants and vowels.

Peripheral Auditory Activity

Once the acoustic signal reaches the ear, it is transduced in a series of steps from pressure variation in the air into the firing of neurons. The varying air pressure first moves the eardrum, then the middle ear bones, the inner ear fluid, the basilar membrane, and finally the hair cells sitting atop the basilar membrane. The stereocilia sticking out of the tops of the hair cells are disturbed by their movements, and that disturbance releases neurotransmitter at the synapse between the hair cells and auditory nerve neurons, causing those neurons to fire. The response to sound is then transmitted via the successive firing of neurons in brainstem nuclei, the inferior colliculus, and the thalamus, up to the auditory cortex. The specific effects of speech sounds on the mechanical stages of transduction, between movement of the eardrum and of the basilar membrane and the hair cells atop of it, have not been studied, but there have been many investigations of how the acoustic properties of speech sounds are represented in the firing patterns of neurons, particularly those in the auditory nerve (Delgutte, 1997). Because these firing patterns are measured by inserting electrodes directly into auditory nerve fibers, none of this work has been done in humans; instead, nonhuman species have been used whose auditory system's characteristics are more or less similar to that of humans, e.g., cats, chinchillas, and guinea pigs. The acoustic properties of the signal are dramatically transformed by their transduction into neuron firing in ways that could profoundly influence the perception of speech sounds, and the probable perceptual effects of these auditory transformations has been used to explain, for example, why vowels are distributed in some ways but not others in the vowel space (Lindblom, 1986) and how stops differing in voice-onset time are categorized (Kluender *et al.,* 1995; cf. Lotto and Kluender, 2002).

Central Nervous System Activity

Activity in the cortical areas of the brain has also been recorded during speech production and perception tasks using electroencephalography (EEG), magnetic encephalography (MEG), positron emission tomography (PET), and functional MRI (fMRI) (Phillips *et al.,* 2000; Phillips, 2001; Gandour *et al.,* 2002, 2003; Guenther, 2003; Guenther and Perkell, 2004). Electroencephalography records voltages at the scalp that reflect parallel currents flowing through large numbers of active neurons, while MEG records the magnetic fields generated by these current flows. The two methods have equally good temporal resolution, on the order of a millisecond or less, but they differ in the location along the cortical surface which produce the strongest signals: changes in voltage are greatest in the direction of current flow, while changes in magnetic field strength are greatest perpendicular to current flow. Electroencephalography is thus most sensitive to radial currents in the gyri, while MEG is most sensitive to such currents in the sulci. The currents that generate these fields are harder to locate in the brain with EEG than MEG, where their location can be determined within a volume equivalent to that of a cube with 5-mm sides. To keep the task of locating the currents mathematically tractable, algorithms assume that just one current is responsible for the observed fields. Because the other measurement techniques often show brain activity in multiple sites, this assumption is a gross simplification.

The other two techniques, PET and fMRI, can locate the site(s) of brain activity with much greater precision because they directly measure changes in blood supply to brain sites as their activity levels changes. As its name indicates, PET measures the emission of positrons. The positrons are emitted by radioactive tracers introduced into the blood by injection or inhalation. As the blood supply increases to a site in the brain when the neurons in that site become more active, more positrons are emitted from more active brain regions. Functional MRI also measures increases in blood flow to more active brain regions, but does so by assessing contrasts between brain sites in the extent to which the iron in hemoglobin is polarized by a strong magnetic field. Because the iron in deoxygenated hemoglobin is much more susceptible to polarization than that in oxygenated hemoglobin, such a contrast will be observed between the predominantly oxygenated blood in the arterioles on one side of a capillary bed and the predominantly deoxygenated blood in the venules on the other side. This contrast will be greater in an active brain region because the arterioles dilate to supply more oxygenated blood to more active regions. The spatial resolution of PET and fMRI is exceptionally good: differences in the numbers of positrons emitted or in magnetic contrast can be detected between volumes as small as a cube with

1 mm sides. However, temporal resolution is relatively poor because changes in blood flow take place over seconds rather than milliseconds.

Speech Synthesis

The first working speech synthesizer was built by Wolfgang von Kempelen. As described in his book, *Mechanismus der menschlichen Sprache nebst Beschreibung einer sprechenden Maschine* (1791), this mechanical device used a bellows to drive air either past an oscillating reed to simulate voiced sounds or through a shunt into the mouth chamber to simulate voiceless ones. The airflow was directed through whistles to simulate different fricative noises. Resonance frequencies were manipulated by covering the opening of the mouth or nose. After some practice, it was possible to produce quite intelligible words and even short sentences with this device.

However, the naturalness of the speech produced by this device cannot compare to that of the many instances of synthetic speech that most of us encounter regularly in the course of our daily lives, for example, when calling directory assistance, booking flights, or listening to announcements in train stations and airports. This naturalness derives only in part from the extensive use of speech synthesizers in laboratory studies of speech perception; the rest comes from a technique widely used in the construction of commercial speech synthesizers: concatenation of demisyllables. (The following account of demisyllable synthesis is greatly simplified.) Naturally produced syllables are divided into initial and final demisyllables. Initial demisyllables consist of the syllable's initial consonant(s), if any, and the transition into the vowel, and final demisyllables consist of the transition from the vowel into the syllable's final consonant(s), if any. Because the transitions between consonants and vowels contain much of the contextual variation between adjacent segments, this variation is encoded directly in the demisyllables and need not be implemented through adjustments of individual consonants and vowels. This solution to the fluent and intelligible synthesis of segment strings has permitted the developers of commercial speech synthesizers to concentrate their research efforts on the other property that bolsters naturalness and intelligibility: prosody. For this reason, demisyllable synthesis is also often used in laboratory studies of the perception of prosody.

Otherwise, most of the synthesizers used in the laboratory are of the type called 'terminal analogue' synthesizers, so called because the user directly manipulates the acoustic properties of the sound source and filter rather than the articulations that determine those acoustic properties. Early synthesizers of this type, e.g., the Pattern Playback developed by Frank Cooper at Haskins Laboratories and the OVE developed by Gunnar Fant at the Royal Institute of Technology in Stockholm, pass an electronic sound source through a filter consisting of parallel resonating circuits. The other, less commonly used method is so-called 'transmission line analogue' synthesis, in which the filter consists of a cascading series of circuits, each representing the cross-sectional area of a short section of the vocal tract. A transmission line analogue synthesizer is thus the first step toward true articulatory synthesis, in which articulators' movements and positions are simulated and their effects on the vocal tract's cross-sectional area are input to the synthesizer for transformation into resonance frequencies and bandwidths. In the last three decades, computer software has replaced electronics in both kinds of synthesizers.

Speech Recognition

If the goal of speech synthesis is to produce an utterance that sounds so natural that the listener thinks it was produced by a human rather than a machine, the goal of speech recognition is recognize the content of naturally produced speech with the same speed and skill that a human can. Successful speech recognition must overcome three very difficult challenges: different speakers' pronunciations of the same utterances differ from one another, words' pronunciations differ between contexts, and speech is often produced in the context of other speech and environmental noises. So far, no speech recognizer has successfully met all three challenges at once.

The vast majority of approaches to speech recognition use statistical pattern recognition and rely on 'hidden Markov models' (HMMs). In a Markov process, the current state depends only on previous states, and the probabilities of a transition from any previous state to the current state are fixed. The Markov model is 'hidden' when the observable states correspond to a hidden Markov process. In the application of HMMs to speech recognition, the observable states are a particular pronunciation of a word, usually encoded as a sequence of spectra, and the hidden Markov process is the speaker's intended sequence of phonemes or articulations for that word. A word is recognized based on the fixed probability that a particular observable state will occur given a particular state in the hidden model.

In principle, speech recognizers could be very useful in laboratory phonetics for explicit and precise testing of hypotheses about humans' recognition of phonemes and words. In practice, however, they have been little used for this purpose, even though phoneme and word recognition by humans is widely

acknowledged to be in part a probabilistic mapping of observable states to hidden processes, i.e., from pronunciations to stored forms. There are two reasons why such recognizers have not been used much to model humans' recognition behavior. First, phoneme and word recognition is only in part a probabilistic mapping, and there is considerable evidence that many of the hidden processes are not Markovian – they are categorical rather than probabilistic. Second, the acoustic properties that are known to influence perception are various in nature and are not easily represented as a series of spectra or any other single acoustic form.

Persistent Issues and Widely Used Experimental Techniques

In this section, a number of central problems in phonetics are described that have remained unsolved, along with the experimental techniques phoneticians have used in wrestling with them, to convey some idea of what kind of work one might see going on in present-day phonetics laboratories. The list is necessarily incomplete, but it nonetheless covers a wide range of problems.

Analysis by Synthesis

The field phonetician was described above as listening carefully to an utterance and then trying to imitate it as closely as possible. Because the phonetician knows, with some precision, how he or she pronounced the imitation, then he or she can infer that the original utterance was pronounced in the same way if the speaker accepts the imitation as accurate. A great deal of research by laboratory phoneticians uses an entirely parallel technique, called 'analysis by synthesis,' where a particular phonetic phenomenon is analyzed by attempting to synthesize it accurately. Because the workings of the synthesizer are known at least as precisely as the phonetician knows the workings of his or her vocal tract, the laboratory phonetician can also infer that the same processes produced the observed phonetic phenomenon if it is synthesized accurately. The motor theory of speech perception (Liberman *et al.*, 1967; Liberman and Mattingly, 1985, 1989) takes this approach to the point of treating it as the means that human listeners use to recognize speech sounds; they do so by simulating the articulations that would produce the acoustic properties they hear.

Another, more general use of analysis by synthesis exists in the development of quantitative models of the transformation from articulations to acoustics. The most influential of these analysis-by-synthesis models has been that developed in Fant's *An acoustic theory of speech production* (1960). Fant used radiographs of a Russian speaker pronouncing the vowels and consonants of this language to determine what the articulatory targets were for each of these sounds. The target for each sound was represented as an 'area function,' which specifies how much the oral cavity is expanded or constricted at each of a number of closely spaced locations along its length. These area functions were then put into a transmission line analogue speech synthesizer, which produced the resonance frequencies corresponding to that area function. This is analysis by synthesis, because the correctness of the articulatory configuration represented by an area function is tested by how closely the synthesizer's acoustic output matches the acoustics of the original sound. This model and its descendants (see Stevens, 1998, particularly) have since been very widely used in investigating the transformation of articulations into acoustics.

Perturbation of Articulations and Feedback

How can one begin to understand what speakers are doing when speaking fluently? Multiple articulators move continuously, tracing intricate patterns. A common answer has been to perturb the articulations in some way and observe the result. For example, bite blocks have been inserted between the teeth to remove the contribution of jaw movement to raising the tongue, and loads have been added to the lower lip to inhibit its raising toward the upper lip in closing the lips (Lindblom *et al.,* 1979; Kelso and Tuller, 1983; Abbs *et al.,* 1984). So long as the perturbation is not too extreme, speakers immediately and successfully shrug it off, i.e., they raise the tongue more independently of the jaw when the bite block prevents them from raising the jaw, and they exert more force to raise the lower lip to overcome the load and lower the upper lip more. These findings not only show what speakers' articulatory goals are but also show that they are capable of reaching those goals in more than one way, each of which is 'motor equivalent' to the others (Kelso *et al.,* 1984; Guenther, 1994). In retrospect, this outcome is unsurprising, because speakers must be capable of reaching a particular articulatory target in the context of any of the articulatory positions that can naturally precede and follow that target.

Recently, a number of studies have examined the effect of perturbing the acoustic feedback that a speaker receives about the sounds he or she has produced (Houde and Jordan, 1998; Jones and Munhall, 2002, 2003); the perturbations are introduced gradually in small, undetectable steps. Speakers adjust their pronunciations to compensate for these perturbations in feedback, just as they do to

perturbations of their articulations. A striking finding in these studies is that compensatory shifts in articulations persist for a considerable time after the perturbations cease – typically many minutes.

Invariance and Variability

Perhaps the most fundamental problem in laboratory phonetics is variability. At some level of description, both the speaker's articulatory goal in pronouncing a particular speech sound and the listener's perception of that pronunciation do not vary with context. For example, the speaker intends to pronounce the same consonant at the beginnings of the words *key* and *coo*, the voiceless velar stop /k/, and the listener perceives that speaker has done so. However, /k/'s pronunciations differ enormously between these two contexts, being far more forward on the palate before the front vowel in *key* than before the back vowel in *coo*, and the acoustics differ comparably. This example of articulatory and acoustic variability is by no means extreme, but instead typical. This example also shows that the targets speakers reach, as well as the paths their articulators follow to reach them, vary enormously if systematically as a function of a potentially very large number of flanking articulatory positions. Each of these many articulatory variants produces a corresponding acoustic variant, which the listener must somehow decode. If the speaker is trying to pronounce one of a small number of different phonemes (or distinctive feature values), which the listener is in turn trying to recognize, then this variation is a serious learning challenge to both of them. Skilled speakers have learned which combinations of articulatory movements are motor equivalent and choose the appropriate one for the current context, but how did they learn to do so? Similarly, skilled listeners have learned how context alters the intended phoneme's acoustics and adjust their criteria for evaluating those acoustics appropriately – but again, how have they learned to do so?

A number of attempts have been made to show that both articulations and acoustics are less variable than they appear. For example, when speakers pronounce vowels and consonants such as [u, ʃ] that require simultaneous articulations with the tongue and the lips, they produce less extreme lingual articulations when the labial articulation is more extreme and vice versa (Perkell *et al.,* 1993, 1998; see also Guenther *et al.,* 1999 for evidence of articulatory tradeoffs in the pronunciation of [ɹ]). This tradeoff is apparently another instance of motor equivalence, but involving entirely independent articulators. What is equivalent, moreover, is the length of the resonating cavity between the two articulations, or perhaps the acoustic product of that resonating cavity, both of which vary far less than they would if these articulations did not trade off in this way. Articulatory phonology (Browman and Goldstein, 1989, 1990, 1992) embodies a more general attempt to show that articulations are actually invariant. This theory proposes that the invariant components of utterances are not abstract feature values but instead concrete gestures that specify the location and degree of constrictions in the vocal tract. Gestures overlap with one another, so a separate component – the 'task dynamics' – is needed in the model to resolve the potentially competing demands of simultaneous gestures (Saltzman and Munhall, 1989). The actual realization of a gesture varies depending on what other gestures it overlaps with, but the gesture's presence in the utterance remains unchanged.

Acoustic invariance has been found for place of articulation in the spectral distribution of energy in the noise bursts that occur when a stop consonant is released (Stevens and Blumstein, 1978; Blumstein and Stevens, 1979, 1980) and in the relationship between the second formant's value at the onset of the transition into a vowel and its value at the vowel's steady state (Sussman *et al.,* 1991). Both measures are acoustically invariant enough that a phonetician can use them to identify the place of articulation correctly in the vast majority of tokens. However, the evidence that listeners use these acoustic invariants is weak: when the information about place of articulation in the noise burst spectra competes with that in formant transitions, which are variable with context rather than invariant, the formant transitions nearly always determine the perceived place (Walley and Carrell, 1983; Smits *et al.,* 1996a, 1996b), and the third formant must begin at a frequency specific to a particular place of articulation, if listeners are to use the relationship between the second formant's frequencies at the transition onset and vowel steady state to identify that place of articulation (Fruchter and Sussman, 1997).

Segmentation and Coarticulation

In the earlier discussion of the acoustic analysis of the phrase 'laboratory phonetics,' both the waveform and the spectrogram were divided up into a number of acoustic segments using more or less obvious visible discontinuities in the displays that corresponded to more or less dramatic acoustic discontinuities in the signal. These segments were, moreover, labeled with symbols that represent the phonemes of which the speaker's message is composed, annotated in some instances to show details of their pronunciation, e.g., the superscript ^{h} that indicates that the /t/ in 'laboratory' is aspirated. However, it is well known that the articulations that realize these phonemes

do not begin and end within the acoustic segments labeled with those symbols. Instead, they begin before (sometimes well before) those segments and end after (sometimes long after) them, i.e., they begin and end within acoustic segments that are labeled with adjacent symbols and that realize adjacent phonemes. Because the articulations of adjacent phonemes overlap or coarticulate with one another, sometimes quite extensively, it is not possible to attribute the acoustic properties of any acoustic segment to a single phoneme's pronunciation. This coarticulation and its acoustic effects are among the principal if systematic sources of the variation in speech; others are differences between speakers and differences between faster and slower speech or between hypo- and hyperarticulated speech.

Categorization and Discrimination

In the vast majority of studies of speech perception, listeners are presented with members of a continuum that differ incrementally from one extreme to another in some acoustic property that is hypothesized to distinguish one phoneme or distinctive feature value from another. The listeners' task is to assign the stimuli to one or the other of a small number of categories, which usually correspond to minimally contrasting phonemes in their native language. For example, listeners might be presented with a series of stops increasing incrementally in the onset frequency of the second formant (F2) and have to categorize them as 'b' versus 'g' (Figure 11a). Figure 12a shows typical results: listeners identify the stimuli at one end of the continuum (stimuli 0 to 1) as one category, here 'b,' and those at the other end (5 to 6) as the other, 'g,' and they cross over from hearing one category to hearing the other within a relatively narrow portion of the continuum (2 to 4). In addition to this identification task, listeners may also be presented with pairs of identical or adjacent stimuli from the continuum and asked whether they are the same or different to measure the extent to which they can discriminate stimuli that differ by a small amount in their values for the acoustic property. If the stimuli are categorized in the way they are in Figure 12a, then discrimination is typically poor (but unlike in the figure detectably above chance) for stimuli that are labeled the same and much better for those that are labeled differently (Figure 12b). Dimensions like F2 onset frequency, where the listeners' ability to discriminate two stimuli depends on the extent to which they label them differently, are described as being 'categorically' perceived.

A quite different pattern of categorization and discrimination responses is obtained for dimensions of simple nonspeech sounds, such as the frequency of a tone. When asked to categorize tones differing in frequency as having 'low' versus 'high' pitch, listeners' 'low' judgments decrease gradually rather than abruptly with frequency (Figure 12c), and when asked to discriminate identical or adjacent stimuli, they are more or less equally good across the range (Figure 12d). This pattern of results is described as 'continuous' perception.

Discrimination performance declines modestly at the high end of the continuum for continua such as the pitch of a tone because it is more difficult to distinguish sounds that differ by a fixed amount when both of their values are high than when they are both low. This effect of stimulus level on discrimination is so general that it has been codified as Weber's Law: the just noticeable difference (jnd) between two stimuli is proportional to the mean of their values for the dimension being discriminated. Discrimination is quite good for the stimulus pairs at the low end of the continuum in Figure 12d, because the difference in pitch between them is much greater than a jnd, but discrimination declines for pairs toward the high end, because this fixed difference gets closer and closer to the jnd. Weber's Law effects are such a general finding in discrimination tasks that failing to find then in the discrimination of categorically perceived dimensions such as F2 onset frequency is particularly surprising.

Even more surprising is the finding that F2 onset frequency is continuously rather than categorically perceived when the F2 transition is presented separately from the rest of the syllable (see the spectrograms of such stimuli in Figure 11b; see also Mattingly *et al.,* 1971). When presented alone, the F2 transition no longer conveys information about place of articulation; indeed, it no longer sounds like speech but instead like a nonspeech chirp or quack. This finding led to the proposal that categorical perception is a special property of humans' responses to acoustic dimensions that distinguish speech sounds from one another (Liberman *et al.,* 1967; Liberman and Mattingly, 1989). The proponents of this view are also the proponents of the motor theory of speech perception. Speech sounds are the only sounds that are perceived categorically because only they are perceived by simulating their production, and they are produced with categorically distinct and largely invariant articulations. However, nonhuman listeners perceive speech dimensions as categorically as humans do (Kuhl and Miller, 1978), and stimuli whose acoustic properties vary like analogous speech sounds but that are not recognized as speech – so-called 'nonspeech analogues' – can be perceived as categorically as speech (Pisoni, 1977; Parker *et al.,* 1986). Moreover, the articulation of particular speech sounds is

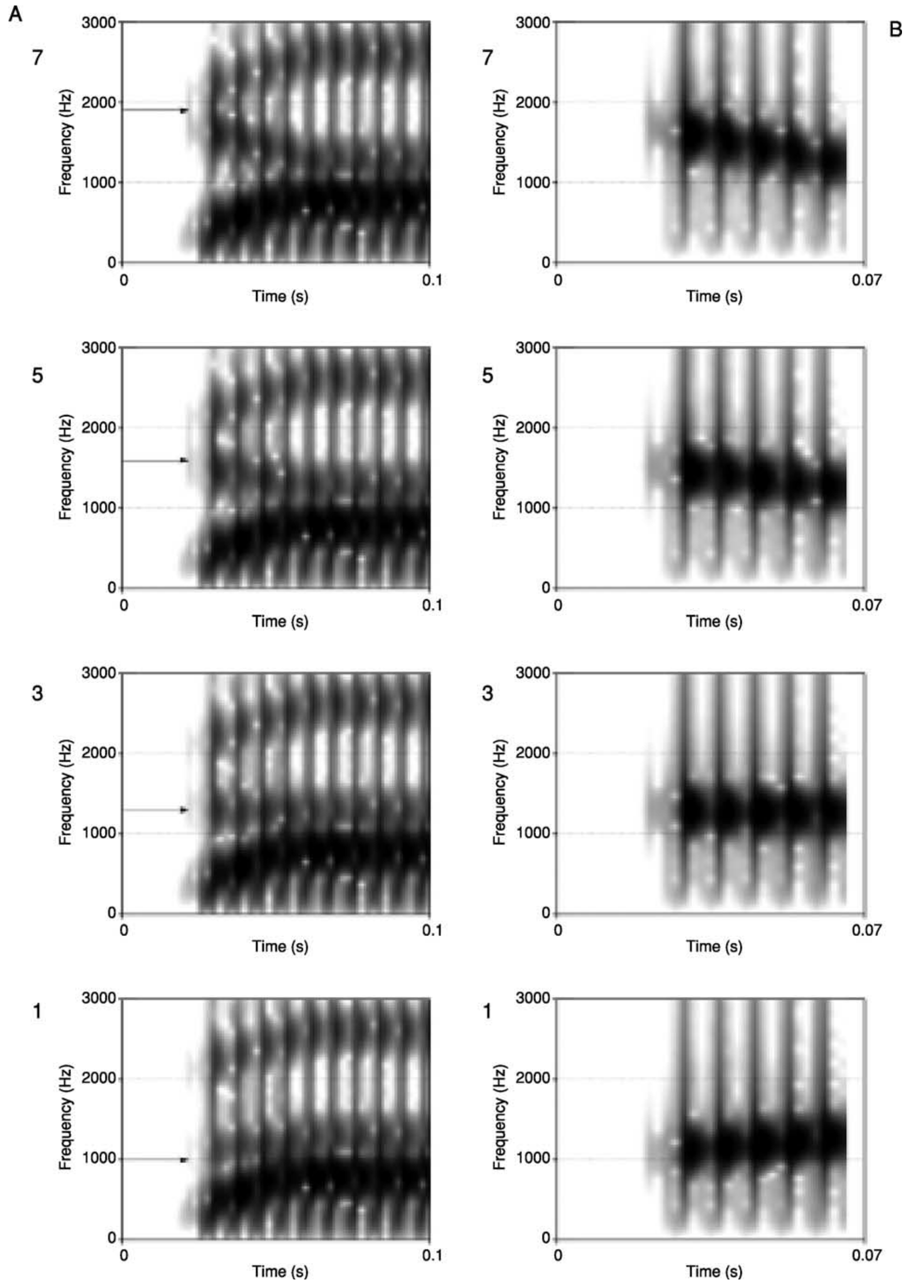


Figure 11 (A) Spectrograms of the first 100 ms of steps 1 (*bottom*), 3, 5, and 7 (*top*) in a seven-step [ba] to [ga] continuum synthesized by varying only the onset frequency of the second formant (F2). The arrows point to the onsets of F2 at 1000 (*bottom*), 1300, 1600, and 1900 Hz. (B) Spectrograms of just the F2 transitions alone.

not invariant with context, even at the level of muscle activity (MacNeilage, 1970). These findings have led others to propose that speech sounds are perceived using general auditory mechanisms rather than speech-specific ones (Diehl and Kluender, 1989).

This debate remains unresolved and now extends to evidence from studies of trading relations, i.e., the change in the perceptual value of one acoustic property by the value of another independently varied property (Repp, 1982; Parker *et al.,* 1986) and of duplex perception, i.e., simultaneous speech and non-speech percepts when the acoustic properties of a syllable are divided between the two ears (Liberman and Mattingly, 1985, 1989; Xu *et al.,* 1997; cf. Fowler and Rosenblum, 1991). The debate has also been extended theoretically with the development of the direct realist theory of speech perception (Fowler, 1986), which claims that speech sounds are perceived

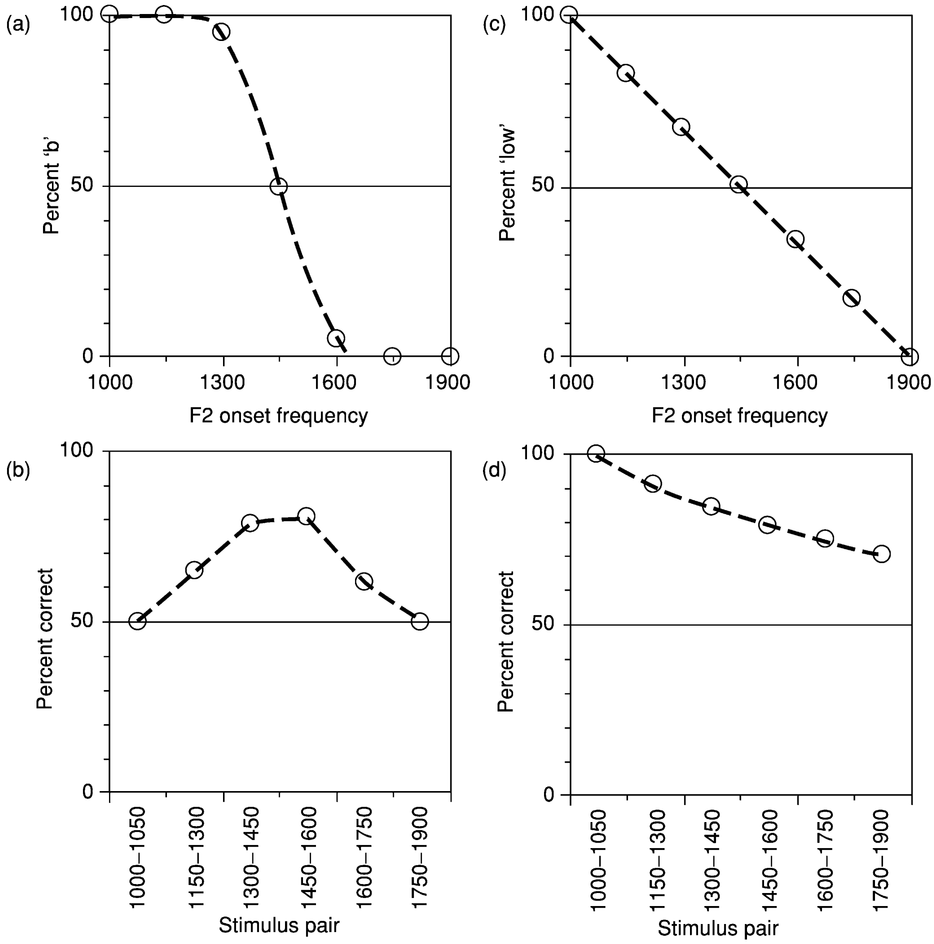


Figure 12 (a) Percent 'b' responses for categorical identification of the F2 onset continuum in Figure 11A, in judgments of place of articulation, and (b) percent correct in discrimination of adjacent stimuli from the continuum depicted in Figure 11A, – chance is 50% correct. (c) Percent 'low' responses to the frequency of a pure tone or the onset of the isolated F2 continuum depicted in Figure 11B, and (d) percent correct in discrimination of adjacent stimuli from the continuum depicted in Figure 11B. (Hypothetical data.)

no differently than other sounds – or for that matter, no differently than stimuli in other modalities: all are perceived as the events that produced them. In direct realism, speech sounds are perceived as the articulations that produce them, but unlike in the motor theory of speech perception as the actual articulations in the mouth of the speaker rather than as simulations of those articulations in the mind of the perceiver.

Phonetic Explanation and Universal Versus Language-Specific Phonetics

Many of the results of phonetic research in the laboratory have been applied very fruitfully in explaining why particular phonological patterns and sound changes recur across languages (Ohala, 1993, 1996). This application has been so fruitful because in the absence of pathology, speakers' vocal tracts and listeners' auditory apparatus work much the same way regardless of what language(s) they speak. These similarities are great enough that speakers' and listeners' phonetic behaviors are constrained in very similar ways across languages. Even so, there is also considerable evidence that speakers of different languages do not pronounce what would appear to be the same phonological categories identically, and these differences, too, may contribute to explanations of sound patterns (Keating *et al.*, 1983; Kingston and Diehl, 1994; Ladefoged and Maddieson, 1996). Until the phonetics of a much greater number of languages are thoroughly documented through careful instrumental analyses of the kinds described in this entry, it will not be possible to determine what the limits are on language-specific phonetic behavior or how far universal constraints imposed by the working of the speaking and listening apparatus extend.

Influence of Types of Linguistic Knowledge on Speech Sound Perception

Finally, in the last two decades there have been many studies of the ways in which listeners' linguistic knowledge affects their perception of speech sounds. Studies of infant listeners have shown that early in the first year of life they are capable of discriminating any sounds that a language might use to distinguish one message from another, but as they approach their first birthday, they gradually lose the ability to discriminate any sounds other than those that contrast in the ambient language (Werker and Tees, 1984; Werker and Lalonde, 1988). Infants show evidence of knowing something about the prosody of the ambient language even earlier: neonates no more than 48 hours old are capable of discriminating their mother's language (not necessarily spoken by her) from a rhythmically different language (Mehler *et al.,* 1988; Nazzi *et al.,* 1998). This ability probably arises from the neonates' having heard the rhythm of their mothers' speech transmitted through her body into the uterus, although premature neonates born as early as 7 months' gestation can also make this distinction, despite having just developed a functioning auditory apparatus (Dehaene-Lambertz, 1998). The effects of linguistic experience increasingly constrain both the speaker's ability to pronounce the sounds of another language and the listener's ability to identify them as the person matures. Even by the time a person reaches his or her fifth birthday, he or she may no longer be able to learn to pronounce or perceive the sounds of another language natively (Pallier *et al.,* 1997).

Mature listeners' perception is not only prevented from distinguishing foreign speech sounds that do not contrast in their native language, but their perception of their native language is also profoundly constrained by the kinds of knowledge they have acquired about it, of whether a particular string of sounds is a word (Ganong, 1980; Pitt and Samuel, 1993), a frequent sound sequence (Pitt and McQueen, 1998), a phonotactically legal sequence (Massaro and Cohen, 1983; Moreton, 2002), or is a word with many versus few phonologically similar neighbors (Vitevitch and Luce, 1998, 1999). The effects of this linguistic knowledge can on occasion be strong enough to produce perceptual illusions (Dupoux *et al.,* 1999) that are even reflected in the brain's earliest detectable response to speech (Dehaene-Lambertz *et al.,* 2000). There is nonetheless considerable debate among students of these effects about whether and how the initial evaluation of the incoming signal is influenced by different kinds of linguistic knowledge, with some arguing that linguistic knowledge affects processing from the outset (Elman and McClelland, 1988) and others arguing instead that linguistic knowledge only biases later decisions about what category the output of initial processing should be assigned to (Norris *et al.,* 2000; Kingston, in press).

Laboratory Speech and Synthetic Speech

Most research in laboratory phonetics uses very carefully designed materials elicited from speakers with a script to study production and equally carefully synthesized materials to study perception. Because much of this research is interested only in how a single property of the utterance is pronounced or perceived, the materials are often not even composed of words, e.g., the production or perception of voicing in intervocalic stops often uses the utterances [aba] and [apa]. Speech elicited in this way is typically very carefully articulated, and as a consequence contrasting sounds differ by large amounts in many phonetic properties. Because the phonetic differences are exhibited so extremely and profusely, this so-called 'laboratory speech' is most useful for studying how minimally contrasting sounds differ from one another. In sharp contrast, no acoustic properties other than those whose perceptual values are of interest differ in the synthetic stimuli used to study perception. This sparseness is most useful for pinpointing how an acoustic property contributes to recognizing that one category has occurred rather than another. Both laboratory speech and synthetic stimuli idealize the phonetics of an utterance and may be quite unrealistic. After all, speakers typically do not declaim their utterances with the distinctness of laboratory speech, and speech sounds in utterances that listeners typically hear differ in many acoustic properties from minimally different sounds. In an effort to study typical speech production and perception, phoneticians have begun to turn to more naturalistically pronounced speech in studying production (Docherty *et al.,* in press) and in studying perception to stimuli that differ in more than one acoustic property at a time (Kingston and Macmillan, 1995; Kingston *et al.,* 1997; Macmillan *et al.,* 1999).

Laboratory Phonology

This entry has described the problems that phoneticians study in the laboratory and the techniques and instruments they use to study them. The purpose of these studies is to discover the systematic patterns in speakers' and listeners' phonetic behavior. But, as acknowledged in the earlier discussion of speech recognition, these patterns are in large part the observable states that correspond to hidden processes, i.e., the phonological patterns of the speakers' and listeners'

language. And as also noted, the results of many studies of phonetic behavior in the laboratory have been successfully applied to explaining recurrent phonological patterns. In an effort to stimulate more investigations of the relationship or interface between phonetics and phonology, the first Laboratory Phonology Conference was held in 1987 at Ohio State University, and the ninth conference was held at the University of Illinois, Urbana-Champaign, in the summer of 2004. The proceedings from the first seven meetings have been published, and the publication of those from the eighth is imminent. As described in the introduction to the first volume (Beckman and Kingston, 1990), the purpose of the conference and proceedings is to change the research practice of both phoneticians and phonologists: to induce phoneticians to test hypotheses about phonetic behavior arising from explicit phonological analyses, and to induce phonologists to recognize that acceptance of a phonological analysis depends on experimentally confirming its predictions about phonetic behavior.

See also: Abercrombie, David (1909–1992); Delattre, Pierre (1903–1969); Evoked Potentials; Evolution of Phonetics and Phonology; Experimental and Instrumental Phonetics: History; Experimental Phonology; Fant, (Carl) Gunnar (Michael) (b. 1919); Fischer-Jørgensen, Eli (b. 1911); fMRI Studies of Language; Fry, Dennis Butler (1907–1983); Gimson, Alfred Charles (1917–1985); Helmholtz, Hermann Ludwig Ferdinand von (1821–1894); Imaging and Measurement of the Vocal Tract; Jones, Daniel (1881–1967); Kempelen, Wolfgang von (1734–1804); Klatt, Dennis H. (1938–1988); Ladefoged, Peter (b. 1925); Lehiste, Ilse (b. 1922); Magnetoencephalography; Malmberg, Bertil (1913–1994); Phonetics, Acoustic; Phonetics, Articulatory; Phonetics: Overview; Rousselot, Pierre Jean, Abbé (1846–1924); Sound Change; Speech Aerodynamics; Speech Perception; Speech Production; Speech Recognition, Audio-Visual; Speech Recognition, Automatic: History; Speech Recognition: Statistical Methods; Speech Synthesis; States of the Glottis; Sweet, Henry (1845–1912).

Bibliography

Abbs J H & Gracco V L (1984). 'Control of complex motor gestures: Orofacial muscle responses to load perturbations of lip during speech.' *Journal of Neurophysiology 51,* 705–723.

Abbs J H, Gracco V L & Cole K J (1984). 'Control of multimovement coordination: Sensorimotor mechanisms in speech motor programming.' *Journal of Motor Behavior 16,* 195–232.

Beckman M E & Kingston J (1990). 'Introduction.' In Kingston J & Beckman M E (eds.) *Papers in laboratory phonology I: Between the grammar and physics of speech.* Cambridge, UK: Cambridge University Press. 1–16.

Blumstein S E & Stevens K N (1979). 'Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants.' *Journal of the Acoustical Society of America 66,* 1001–1017.

Blumstein S E & Stevens K N (1980). 'Perceptual invariance and onset spectra for stop consonants in different vowel environments.' *Journal of the Acoustical Society of America 67,* 648–662.

Browman C P & Goldstein L (1989). 'Articulatory gestures as phonological units.' *Phonology 6,* 201–251.

Browman C P & Goldstein L (1990). 'Tiers in articulatory phonology, with some implications for casual speech.' In Kingston J & Beckman M E (eds.) *Papers in laboratory phonology I: Between the grammar and physics of speech.* Cambridge, UK: Cambridge University Press. 341–376.

Browman C P & Goldstein L (1992). 'Articulatory phonology: an overview.' *Phonetica 49,* 155–180.

Byrd D, Browman C P, Goldstein L & Honorof D (1999). 'Magnetometer and X-ray microbeam comparison.' In Ohala J J, Hasegawa Y, Ohala M, Granville D & Bailey A C (eds.) *Proceedings of the 14th International Congress of Phonetic Sciences.* San Francisco. 627–630.

Dart S N (1991). *Articulatory and acoustic properties of apical and laminal articulations.* Ph.D. diss., Los Angeles: University of California.

Dart S N (1998). 'Comparing French and English coronal consonant articulation.' *Journal of Phonetics 26,* 71–94.

Dehaene-Lambertz G (1998). 'Syllable discrimination by premature neonates with or without subcortical lesion. Biology of the neonate.' *Developmental Neuropsychology 14,* 579–597.

Dehaene-Lambertz G, Dupoux E & Gout A (2000). 'Electrophysiological correlates of phonological processing: A crosslinguistic study.' *Journal of Cognitive Neuroscience 12,* 635–647.

Delgutte B (1997). 'Auditory neural processing of speech.' In Hardcastle W J & Laver J (eds.) *Handbook of phonetic sciences.* Oxford: Blackwell. 507–538.

Demolin D, Metens T & Soquet A (2000). 'Real time MRI and articulatory coordinations in vowels.' Paper presented at the Fifth Seminar on Speech Production: Models and Data, Kloster.

Docherty G J, Foulkes P, Tillotson J & Watt D J L (in press). 'On the scope of phonological learning: issues arising from socially structured variation.' In Best C T, Goldstein L & Whalen D H (eds.) *Laboratory phonology 8.* Berlin: Mouton de Gruyter.

Dupoux E, Kakehi K, Hirose Y, Pallier C & Mehler J (1999). 'Epenthetic vowels in Japanese: A perceptual illusion.' *Journal of Experimental Psychology: Human Perception and Performance 25,* 1568–1578.

Elman J L & McClelland J L (1988). 'Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes.' *Journal of Memory and Language 27,* 143–165.

Eulitz C & Lahiri A (2004). 'Neurobiological evidence for abstract phonological representations in speech recognition.' *Journal of Cognitive Neuroscience 16,* 577–583.

Fant G (1960). *An acoustic theory of speech production.* The Hague: Mouton.

Fowler C A (1986). 'An event approach to the study of speech perception from a direct-realist perspective.' *Journal of Phonetics 14,* 3–28.

Fowler C A & Rosenblum L D (1991). 'Duplex perception: a comparison of monosyllables and slamming doors.' *Journal of Experimental Psychology: Human Perception and Performance 16,* 742–754.

Fruchter D & Sussman H M (1997). 'The perceptual relevance of locus equations.' *Journal of the Acoustical Society of America 102,* 2997–3008.

Gandour J, Wong D, Lowe M, Dzemidzic M, Satthamnuwong N, Tong Y & Li X (2002). 'A crosslinguistic fMRI study of spectral and temporal cues underlying phonological processing.' *Journal of Cognitive Neuroscience 14,* 1076–1087.

Gandour J, Xu Y, Wong D, Dzemidzic M, Lowe M, Li X & Tong Y (2003). 'Neural correlates of segmental and tonal information in speech perception.' *Human Brain Mapping 20,* 185–200.

Ganong W F III (1980). 'Phonetic categorization in auditory word perception.' *Journal of Experimental Psychology: Human Perception and Performance 6,* 110–125.

Gibbon F & Nicolaidis K (1999). 'Palatography.' In Hardcastle W J & Hewlett N (eds.) *Coarticulation in speech production: theory, data, and techniques.* Cambridge, UK: Cambridge University Press. 229–245.

Guenther F H (1994). 'A neural network model of speech acquisition and motor equivalent speech production.' *Biological Cybernetics 72,* 43–53.

Guenther F H (2003). 'Neural control of speech movements.' In Meyer A & Schiller N (eds.) *Phonetics and phonology in language comprehension and production: differences and similarities.* Berlin: Mouton de Gruyter. 209–239.

Guenther F H & Perkell J S (2004). 'A neural model of speech production and its application to studies of the role of auditory feedback in speech.' In Maassen B, Kent R, Peters H, Van Lieshout P & Hulstijn W (eds.) *Speech motor control in normal and disordered speech.* Oxford: Oxford University Press. 29–49.

Guenther F H, Espy-Wilson C Y, Boyce S E, Matthies M L, Zandipour M & Perkell J S (1999). 'Articulatory tradeoffs reduce acoustic variability during American English // production. *Journal of the Acoustical Society of America 105,* 2854–2865.

Hess W (1983). *Pitch determination in speech signals.* Berlin: Springer Verlag.

Houde J F & Jordan M I (1998). 'Sensorimotor adaptation in speech production.' *Science 279,* 1213–1216.

Jones J A & Munhall K G (2002). 'The role of auditory feedback during phonation: studies of Mandarin tone production.' *Journal of Phonetics 30,* 303–320.

Jones J A & Munhall K G (2003). 'Learning to produce speech with an altered vocal tract: the role of auditory feedback.' *Journal of the Acoustical Society of America 113,* 532–543.

Keating P A, Linker W & Huffman M (1983). 'Patterns of allophone distribution for voiced and voiceless stops.' *Journal of Phonetics 11,* 277–290.

Kelso J A S & Tuller B (1983). '"Compensatory articulation" under conditions of reduced afferent information: a dynamic formulation.' *Journal of Speech and Hearing Research 26,* 217–224.

Kelso J A S, Tuller B, Vatikiotis-Bateson E & Fowler C A (1984). 'Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordinative structures.' *Journal of Experimental Psychology: Human Perception and Performance 10,* 812–832.

Kempelen W von (1791). *Mechanismus der menschlichen sprache nebst Beschreibung einer sprechenden Maschine.*

Kingston J (in press). 'Ears to categories: new arguments for autonomy.' In Frota S, Vigario M & Freitas M (eds.) *Proceedings of the First Conference on Phonetics and Phonology in Iberia.* Berlin: Mouton de Gruyter.

Kingston J & Diehl R L (1994). 'Phonetic knowledge.' *Language 70,* 419–454.

Kingston J & Macmillan N A (1995). 'Integrality of nasalization and F1 in vowels in isolation and before oral and nasal consonants: a detection–theoretic application of the Garner paradigm.' *Journal of the Acoustical Society of America 97,* 1261–1285.

Kingston J, Macmillan N A, Dickey L W, Thorburn R & Bartels C (1997). 'Integrality in the perception of tongue root position and voice quality in vowels.' *Journal of the Acoustical Society of America 101,* 1696–1709.

Kluender K R, Lotto A J & Jenison R L (1995). 'Perception of voicing for syllable–initial stops at different intensities: Does synchrony capture signal voiceless stop consonants?' *Journal of the Acoustical Society of America 97,* 2552–2567.

Koenig R (1872). 'Die manometrischen Flammen.' *Annalen der Physik und Chemie 161.*

Kuhl P K & Miller J D (1978). 'Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli.' *Journal of the Acoustical Society of America 63,* 905–917.

Ladefoged P & Maddieson I (1996). *The sounds of the world's languages.* Oxford: Blackwell Publishers.

Liberman A M & Mattingly I G (1985). 'The motor theory of speech perception revised.' *Cognition 21,* 1–36.

Liberman A M & Mattingly I G (1989). 'The specialization for speech perception.' *Science 243,* 489–494.

Liberman A M, Cooper F S, Shankweiler D P & Studdert-Kennedy M (1967). 'Perception of the speech code.' *Psychological Review 74,* 431–461.

Lindblom B (1986). 'Phonetic universals in vowel systems.' In Ohala J J & Jaeger J J (eds.) *Experimental phonology.* New York: Academic Press. 13–44.

Lindblom B, Lubker J & Gay T (1979). 'Formant frequencies of some fixed-mandible vowels and a model of speech motor programming by predictive simulation.' *Journal of Phonetics 7,* 147–161.

Lotto A J & Kluender K R (2002). 'Synchrony capture hypothesis fails to account for effects of amplitude on

voicing perception.' *Journal of the Acoustical Society of America 111,* 1056–1062.
Macmillan N A, Kingston J, Thorburn R, Walsh Dickey L & Bartels C (1999). 'Integrality of nasalization and F_1. II. Basic sensitivity and phonetic labeling measure distinct sensory and decision-rule interactions.' *Journal of the Acoustical Society of America 106,* 2913–2932.
MacNeilage P F (1970). ''Motor control' serial ordering of speech.' *Psychological Review 77,* 182–196.
Markel J D & Gray A H (1976). *Linear prediction of speech.* Berlin: Springer Verlag.
Massaro D & Cohen M (1983). 'Phonological context in speech perception.' *Perception and Psychophysics 34,* 338–348.
Mattingly I G, Liberman A M, Syrdal A K & Halwes T (1971). 'Discrimination in speech and nonspeech modes.' *Cognitive Psychology 2,* 131–157.
Mehler J, Jusczyk P, Lambertz G, Hoalsted N, Bertoncini J & Amiel-Tison C (1988). 'A precursor of language acquisition in young infants.' *Cognition 29,* 144–178.
Moore D M & Berke G S (1988). 'The effect of laryngeal nerve stimulation on phonation: a glottographic study using an *in vivo* canine model.' *Journal of the Acoustical Society of America 83,* 705–715.
Moreton E (2002). 'Structural constraints in the perception of English stop-sonorant clusters.' *Cognition 84,* 55–71.
Munhall K G & Vatikiotis-Bateson E (1998). 'The moving face during speech communication.' In Campbell R, Dodd B & Burnham D (eds.) *Hearing by eye, part 2: The psychology of speech reading and audiovisual speech.* London: Taylor & Francis-Psychology Press. 123–139.
Nadler R D, Abbs J H & Fujimura O (1987). 'Speech Movement Research using the New X-Ray Microbeam System.' In *Proceedings of the 11th International Congress of Phonetic Sciences, Tallinn, Estonia, USSR.* Tallin: Academy of Sciences of the Estonian SSR, Institute of Language and Literature. 221–224.
Narayanan S, Nayak K, Lee S, Sethy A & Byrd D (2004). 'An approach to real-time magnetic resonance imaging for speech production.' *Journal of the Acoustical Society of America 115,* 1771–1776.
Nazzi T, Bertoncini J & Mehler J (1998). 'Language discrimination by newborns: towards an understanding of the role of rhythm.' *Journal of Experimental Psychology Human Perception and Performance 24,* 756–766.
Norris D, McQueen J M & Cutler A (2000). 'Merging information in speech recognition: feedback is never necessary.' *Behavioral and Brain Sciences 23,* 299–370.
Obleser J, Eulitz C & Lahiri A (2004). 'Magnetic brain response mirrors extraction of phonological features from spoken vowels.' *Journal of Cognitive Neuroscience 16,* 31–39.
Ohala J J (1990). 'Respiratory activity in speech.' In Hardcastle W J & Marchal A (eds.) *Speech production and speech modelling.* Dordrecht: Kluwer. 23–53.
Ohala J J (1993). 'The phonetics of sound change.' In Jones C (ed.) *Historical linguistics: problems and perspectives.* London: Longman. 237–278.
Ohala J J (1996). 'The relation between phonetics and phonology.' In Hardcastle W & Laver J (eds.) *Handbook of phonetics.* Oxford: Blackwell. 674–694.
Pallier C, Bosch L & Sebastian-Gallés N (1997). 'A limit on behavioral plasticity in speech perception.' *Cognition 64,* B9–B17.
Parker E M, Diehl R L & Kluender K R (1986). 'Trading relations in speech and nonspeech.' *Perception and Psychophysics 39,* 129–142.
Perkell J S, Cohen M I, Svirsky M A, Matthies M L, Garabieta I & Jackson M T T (1992). 'Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements.' *Journal of the Acoustical Society of America 92,* 3078–3096.
Perkell J S, Matthies M L, Svirsky M A & Jordan M I (1993). 'Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: a pilot "motor equivalence" study.' *Journal of the Acoustical Society of America 93,* 2948–2961.
Perkell J S, Matthies M L & Zandipour M (1998). 'Motor equivalence in the production of /ʃ/. *Journal of the Acoustical Society of America 103,* 3085.
Perlman A L & Titze I R (1988). 'Development of an *in vitro* technique for measuring elastic properties of vocal fold tissue.' *Journal of Speech and Hearing Research 31,* 288–298.
Phillips C (2001). 'Levels of representation in the electrophysiology of speech perception.' *Cognitive Science 25,* 711–731.
Phillips C, Pellathy T, Marantz A, Yellin E, Wexler K, McGinnis M, Poeppel D & Roberts T (2000). 'Auditory cortex accesses phonological categories: an MEG mismatch study.' *Journal of Cognitive Neuroscience 12,* 1038–1055.
Pisoni D B (1977). 'Identification and discrimination of the relative onset time of two component tones: implications for voicing perception in stops.' *Journal of the Acoustical Society of America 61,* 1352–1361.
Pitt M A & McQueen J M (1998). 'Is compensation for coarticulation mediated by the lexicon?' *Journal of Memory and Language 39,* 347–370.
Pitt M A & Samuel A G (1993). 'An empirical and meta-analytic evaluation of the phoneme identification task.' *Journal of Experimental Psychology: Human Perception and Performance 19,* 699–725.
Repp B H (1982). 'Phonetic trading relations and context effects: new experimental evidence for a speech mode of perception.' *Psychological Bulletin 92,* 81–110.
Rousselot J-P (1897–1901/1902–1908). *Principes de Phonétique Expérimentale* (2 vols). Paris: H. Welter.
Saltzman E L & Munhall K G (1989). 'A dynamical approach to gestural patterning in speech production.' *Ecological Psychology 1,* 333–382.
Scripture E W (1902). *The elements of experimental phonetics.* New York: Charles Scribner's Sons.
Smits R, ten Bosch L & Collier R (1996a). 'Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants. I. Perception experiment.' *Journal of the Acoustical Society of America 100,* 3852–3864.

Smits R, ten Bosch L & Collier R (1996b). 'Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants. II. Modeling and evaluation.' *Journal of the Acoustical Society of America 100,* 3865–3881.
Stevens K N (1998). *Acoustic Phonetics.* Cambridge, MA: MIT Press.
Stevens K N & Blumstein S E (1978). 'Invariant cues for place of articulation in stop consonants.' *Journal of the Acoustical Society of America 64,* 1358–1368.
Stone M & Lundberg L (1996). 'Three-dimensional tongue surface shapes of English consonants and vowels.' *Journal of the Acoustical Society of America 99,* 1–10.
Stone M, Shawker T, Talbot L & Rich A (1988). 'Cross-sectional tongue shape during the production of vowels.' *Journal of the Acoustical Society of America 83,* 1586–1596.
Stone M, Faber A, Raphael L J & Shawker T H (1992). 'Cross-sectional tongue shape and linguopalatal contact patterns in [s], [ʃ], and [l].' *Journal of Phonetics 20,* 253–270.
Sussman H M, McCaffrey H A & Matthews S A (1991). 'An investigation of locus equations as a source of relational invariance for stop place categorization.' *Journal of the Acoustical Society of America 90,* 1309–1325.
Titze I R, Jiang J J & Lin E (1997). 'The dynamics of length change in canine vocal folds.' *Journal of Voice 11,* 267–276.
Vitevitch M S & Luce P A (1998). 'When words compete: levels of processing in spoken word recognition.' *Psychological Science 9,* 325–329.
Vitevitch M S & Luce P A (1999). 'Probabilistic phonotactics and neighborhood activation in spoken word recognition.' *Journal of Memory and Language 40,* 374–408.
Walley A C & Carrell T D (1983). 'Onset spectra and formant transitions in the adult's and child's perception of place of articulation in stop consonants.' *Journal of the Acoustical Society of America 73,* 1011–1022.
Werker J F & Lalonde C E (1988). 'Cross-language speech perception: initial capabilities and developmental change.' *Developmental Psychology 24,* 672–683.
Werker J F & Tees R C (1984). 'Cross-language speech perception: evidence for perceptual reorganization during the first year of life.' *Infant Behavior and Development 7,* 49–63.
Xu Y, Liberman A M & Whalen D H (1997). 'On the immediacy of phonetic perception.' *Psychological Science 8,* 358–362.

Labov, William (b. 1927)

D Stein, Heinrich-Heine-Universitaet Duesseldorf, Duesseldorf, Germany

Similar to several other linguists, William Labov started out in the natural sciences, as a chemist, before turning into a linguist. He studied and taught at Columbia University (1961) before becoming professor of linguistics at the University of Pennsylvania in 1970. He is considered the founder of modern sociolinguistics, initiating a paradigm of studying variation and change that has been termed 'correlative,' in that it establishes systematic correlations between social categories like class, age, sex, ethnicity, and style as independent variables, and the frequency of variants of (mostly phonological) variables as dependent variables. Apart from now widely applied concepts of variational analysis such as the 'linguistic variable' or his notion of 'style,' Labov's main early theoretical contribution may be seen in the demonstration of the structuredness of linguistic variation against the backdrop of generative disregard of performance as of less interest for the study of language. Labov is also to be credited with the first major study of urban dialects (those of New York City) and for developing a specific methodology of structured interviews of random samples of groups of speakers, which are taken to be representative both for the members of the whole group and, more implicitly, for the individual. His paradigm-setting methodology has been widely adopted for the analysis of other varieties of English and other languages. It was above all the discovery of age-based differences in the use of variants that led to insights into the micromechanisms of present-day linguistic change (Weinreich *et al.*, 1968) and was then extended into the analysis of disputed cases of historical change, such as the English Great Vowel Shift. Labov's work on variation, although focused on phonological change, was also extended into other areas, such as semantic and morphosyntactic variation. Among his major contributions to the study of diffusion patterns in linguistic change processes is the discovery that women are consistently at the forefront of changes involving social prestige, whereas men lead the way in changes motivated by 'covert prestige.'

Another focus of Labov's work concerns educational issues arising from the study of ethnic urban dialects such as African American Vernacular English, centering on grammatical differences between standard English and vernacular grammar. Labov identified

Smits R, ten Bosch L & Collier R (1996b). 'Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants. II. Modeling and evaluation.' *Journal of the Acoustical Society of America 100,* 3865–3881.
Stevens K N (1998). *Acoustic Phonetics.* Cambridge, MA: MIT Press.
Stevens K N & Blumstein S E (1978). 'Invariant cues for place of articulation in stop consonants.' *Journal of the Acoustical Society of America 64,* 1358–1368.
Stone M & Lundberg L (1996). 'Three-dimensional tongue surface shapes of English consonants and vowels.' *Journal of the Acoustical Society of America 99,* 1–10.
Stone M, Shawker T, Talbot L & Rich A (1988). 'Cross-sectional tongue shape during the production of vowels.' *Journal of the Acoustical Society of America 83,* 1586–1596.
Stone M, Faber A, Raphael L J & Shawker T H (1992). 'Cross-sectional tongue shape and linguopalatal contact patterns in [s], [ʃ], and [l].' *Journal of Phonetics 20,* 253–270.
Sussman H M, McCaffrey H A & Matthews S A (1991). 'An investigation of locus equations as a source of relational invariance for stop place categorization.' *Journal of the Acoustical Society of America 90,* 1309–1325.
Titze I R, Jiang J J & Lin E (1997). 'The dynamics of length change in canine vocal folds.' *Journal of Voice 11,* 267–276.
Vitevitch M S & Luce P A (1998). 'When words compete: levels of processing in spoken word recognition.' *Psychological Science 9,* 325–329.
Vitevitch M S & Luce P A (1999). 'Probabilistic phonotactics and neighborhood activation in spoken word recognition.' *Journal of Memory and Language 40,* 374–408.
Walley A C & Carrell T D (1983). 'Onset spectra and formant transitions in the adult's and child's perception of place of articulation in stop consonants.' *Journal of the Acoustical Society of America 73,* 1011–1022.
Werker J F & Lalonde C E (1988). 'Cross-language speech perception: initial capabilities and developmental change.' *Developmental Psychology 24,* 672–683.
Werker J F & Tees R C (1984). 'Cross-language speech perception: evidence for perceptual reorganization during the first year of life.' *Infant Behavior and Development 7,* 49–63.
Xu Y, Liberman A M & Whalen D H (1997). 'On the immediacy of phonetic perception.' *Psychological Science 8,* 358–362.

Labov, William (b. 1927)

D Stein, Heinrich-Heine-Universitaet Duesseldorf, Duesseldorf, Germany

Similar to several other linguists, William Labov started out in the natural sciences, as a chemist, before turning into a linguist. He studied and taught at Columbia University (1961) before becoming professor of linguistics at the University of Pennsylvania in 1970. He is considered the founder of modern sociolinguistics, initiating a paradigm of studying variation and change that has been termed 'correlative,' in that it establishes systematic correlations between social categories like class, age, sex, ethnicity, and style as independent variables, and the frequency of variants of (mostly phonological) variables as dependent variables. Apart from now widely applied concepts of variational analysis such as the 'linguistic variable' or his notion of 'style,' Labov's main early theoretical contribution may be seen in the demonstration of the structuredness of linguistic variation against the backdrop of generative disregard of performance as of less interest for the study of language. Labov is also to be credited with the first major study of urban dialects (those of New York City) and for developing a specific methodology of structured interviews of random samples of groups of speakers, which are taken to be representative both for the members of the whole group and, more implicitly, for the individual. His paradigm-setting methodology has been widely adopted for the analysis of other varieties of English and other languages. It was above all the discovery of age-based differences in the use of variants that led to insights into the micromechanisms of present-day linguistic change (Weinreich *et al.*, 1968) and was then extended into the analysis of disputed cases of historical change, such as the English Great Vowel Shift. Labov's work on variation, although focused on phonological change, was also extended into other areas, such as semantic and morphosyntactic variation. Among his major contributions to the study of diffusion patterns in linguistic change processes is the discovery that women are consistently at the forefront of changes involving social prestige, whereas men lead the way in changes motivated by 'covert prestige.'

Another focus of Labov's work concerns educational issues arising from the study of ethnic urban dialects such as African American Vernacular English, centering on grammatical differences between standard English and vernacular grammar. Labov identified

local social group identity as a major force in explaining the absence of patterns of convergence toward the prestige standard norm. In addition, he stressed the coexistence of two equally functional but separate systems and debunked the claim that the nonstandard system was deficient in comparison to the standard system. He related the educational failure of these groups to the fact that the educational system did not take into account the difference in the systems and to the existence of a nonstandard value system that existed in opposition to that of the school, as part of local group identity. Labov attributed an even stronger effect on educational failure to the existence of negative attitudes toward nonstandard language.

Labov's more recent work continues the orientation toward urban dialectology and change as embodied in the *Atlas of North American English*, conceptualizing the pattern of phonological change observed in a notion of 'cascading' (i.e., moving in cohortal fashion from the largest city to the respective next-largest ones). The sorting out of the role of the sexes has been one of his most important contributions (cf. his extremely important article published in *Language Variation and Change*.) This attention to the role of gender in change permeates his work.

see also: English, African American Vernacular; Society and Language: Overview; Variation and Language: Overview.

Bibliography

Labov W (1966). *The social stratification of English in New York City.* Washington, DC: Center for Applied Linguistics.

Labov W (1969a). 'Contraction, deletion, and inherent variability of the English copula.' *Language 45,* 715–762.

Labov W (1969b). 'The logic of non-standard English.' In Alatis J (ed.) *Georgetown monograph on languages and linguistics 22.* Washington, DC: Georgetown University Press. 1–44.

Labov W (1970). *The study of non-standard English.* Champaign: National Council of Teachers of English.

Labov W (1972a). *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Labov W (1972b). *Language in the inner city.* Philadelphia: University of Pennsylvania Press.

Labov W (1973a). 'The boundaries of words and their meanings.' In Bailey C-J & Shuy R (eds.) *New ways of analyzing variation in English.* Washington, DC: Georgetown University Press. 340–373.

Labov W (1973b). 'The linguistic consequences of being a lame.' *Language in Society 2,* 81–115.

Labov W (1975). 'On the use of the present to explain the past.' In Heilmann L (ed.) *Proceedings of the 11th International Congress of Linguists.* Bologna: Il Mulino. 825–851.

Labov W (1990a). 'The intersection of sex and social class in the course of linguistic change.' *Language Variation and Change 2,* 205–254.

Labov W (1990b). 'On the adequacy of natural languages I: the development of tense.' In Singler J (ed.) *Pidgin and creole tense-mood-aspect systems.* Amsterdam: Benjamins. 1–58.

Labov W, Ash S & Boberg C (in press). *The atlas of North American English.* Berlin: Mouton de Gruyter.

Labov W & Waletzky J (1967). 'Narrative analysis.' In Helm J (ed.) *Essays on the verbal and visual arts.* Seattle: University of Washington Press. 12–44.

Weinreich U, Labov W & Herzog M (1968). 'Empirical foundations for a theory of language change.' In Lehmann W & Malkiel Y (eds.) *Directions for historical linguistics.* Austin: University of Texas Press. 95–195.

Lacan, Jacques (1901–1981)

M Bednarek, University of Augsburg, Augsburg, Germany

Lacan was born in Paris in 1901, studied medicine there, and gained a Ph.D. in 1932 with a thesis on paranoia. Although he was widely influential across disciplinary boundaries (by virtue of his rereading of Freud through the lens of structural linguistics), within his own scientific community he was regarded as an *enfant terrible* and excluded from national and international psychoanalytic societies. Consequently, he founded his own school of psychoanalysis (*L'École Freudienne de Paris*). Lacan worked and lived in Paris until his death in 1981.

Influenced by Claude Lévi-Strauss's structural anthropology, Lacan adopted several ideas from Ferdinand de Saussure and Roman Jakobson in his rewriting of Sigmund Freud's theory (arguing for a 'return to Freud'), the most important essays of which were published in 1966 in the *Écrits*.

His central hypothesis is that the unconscious is structured like a language, in terms of condensation (metaphor) and displacement (metonymy). He also reinterprets Saussure's conception of the linguistic sign (*signifiant* and *signifié*), arguing that there is no stable meaning, no ultimate *signifié* that the *signifiant*

local social group identity as a major force in explaining the absence of patterns of convergence toward the prestige standard norm. In addition, he stressed the coexistence of two equally functional but separate systems and debunked the claim that the nonstandard system was deficient in comparison to the standard system. He related the educational failure of these groups to the fact that the educational system did not take into account the difference in the systems and to the existence of a nonstandard value system that existed in opposition to that of the school, as part of local group identity. Labov attributed an even stronger effect on educational failure to the existence of negative attitudes toward nonstandard language.

Labov's more recent work continues the orientation toward urban dialectology and change as embodied in the *Atlas of North American English*, conceptualizing the pattern of phonological change observed in a notion of 'cascading' (i.e., moving in cohortal fashion from the largest city to the respective next-largest ones). The sorting out of the role of the sexes has been one of his most important contributions (cf. his extremely important article published in *Language Variation and Change*.) This attention to the role of gender in change permeates his work.

see also: English, African American Vernacular; Society and Language: Overview; Variation and Language: Overview.

Bibliography

Labov W (1966). *The social stratification of English in New York City.* Washington, DC: Center for Applied Linguistics.

Labov W (1969a). 'Contraction, deletion, and inherent variability of the English copula.' *Language 45,* 715–762.

Labov W (1969b). 'The logic of non-standard English.' In Alatis J (ed.) *Georgetown monograph on languages and linguistics 22.* Washington, DC: Georgetown University Press. 1–44.

Labov W (1970). *The study of non-standard English.* Champaign: National Council of Teachers of English.

Labov W (1972a). *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Labov W (1972b). *Language in the inner city.* Philadelphia: University of Pennsylvania Press.

Labov W (1973a). 'The boundaries of words and their meanings.' In Bailey C-J & Shuy R (eds.) *New ways of analyzing variation in English.* Washington, DC: Georgetown University Press. 340–373.

Labov W (1973b). 'The linguistic consequences of being a lame.' *Language in Society 2,* 81–115.

Labov W (1975). 'On the use of the present to explain the past.' In Heilmann L (ed.) *Proceedings of the 11th International Congress of Linguists.* Bologna: Il Mulino. 825–851.

Labov W (1990a). 'The intersection of sex and social class in the course of linguistic change.' *Language Variation and Change 2,* 205–254.

Labov W (1990b). 'On the adequacy of natural languages I: the development of tense.' In Singler J (ed.) *Pidgin and creole tense-mood-aspect systems.* Amsterdam: Benjamins. 1–58.

Labov W, Ash S & Boberg C (in press). *The atlas of North American English.* Berlin: Mouton de Gruyter.

Labov W & Waletzky J (1967). 'Narrative analysis.' In Helm J (ed.) *Essays on the verbal and visual arts.* Seattle: University of Washington Press. 12–44.

Weinreich U, Labov W & Herzog M (1968). 'Empirical foundations for a theory of language change.' In Lehmann W & Malkiel Y (eds.) *Directions for historical linguistics.* Austin: University of Texas Press. 95–195.

Lacan, Jacques (1901–1981)

M Bednarek, University of Augsburg, Augsburg, Germany

Lacan was born in Paris in 1901, studied medicine there, and gained a Ph.D. in 1932 with a thesis on paranoia. Although he was widely influential across disciplinary boundaries (by virtue of his rereading of Freud through the lens of structural linguistics), within his own scientific community he was regarded as an *enfant terrible* and excluded from national and international psychoanalytic societies. Consequently, he founded his own school of psychoanalysis (*L'École Freudienne de Paris*). Lacan worked and lived in Paris until his death in 1981.

Influenced by Claude Lévi-Strauss's structural anthropology, Lacan adopted several ideas from Ferdinand de Saussure and Roman Jakobson in his rewriting of Sigmund Freud's theory (arguing for a 'return to Freud'), the most important essays of which were published in 1966 in the *Écrits*.

His central hypothesis is that the unconscious is structured like a language, in terms of condensation (metaphor) and displacement (metonymy). He also reinterprets Saussure's conception of the linguistic sign (*signifiant* and *signifié*), arguing that there is no stable meaning, no ultimate *signifié* that the *signifiant*

refers to. Without the existence of stable *signifiés*, the unconscious is the place of a sliding chain of *signifiants*.

Another of Lacan's well-known concepts is that of the 'mirror stage.' He suggests that the human infant is inherently fragmented and incomplete, a fact which is masked by the child's identification with its mirror image as a unified 'gestalt.' This identification process results in the formation of the self (*moi*) and causes the child's alienation from itself, the split subject. Once the child learns to use a language, it enters the Symbolic order, one of three domains of experience established by Lacan (the other two domains are the Real, where no language exists, and the Imaginary, the order of illusions). This symbolic order is regarded as the 'law' that governs the whole of man's sociocultural existence.

According to Lacan, language is also essential for any psychoanalytic exchange whose aim is to help the subject arrive at *parole pleine* ('full speech') via *parole vide* ('empty speech'), a necessary detour that enables the access of the true subject (*je*).

Although Lacan's style is very complex and idiosyncratic (and defies reading and interpretation to some extent), he had great influence, particularly on philosophy, sociology, cultural studies, and feminist theory, but also on interdisciplinary linguistic approaches such as Critical Discourse Analysis.

See also: Freud, Sigmund (1856–1939); Jakobson, Roman (1896–1982); Jakobson, Roman: Theory of the Sign; Levi-Strauss, Claude (b. 1908); Saussure, Ferdinand (-Mongin) de (1857–1913); Structuralism.

Bibliography

Dylan E (1996). *An introductory dictionary of Lacanian psychoanalysis*. London: Routledge.

Felman S (1987). *Jacques Lacan and the adventure of insight. Psychoanalysis in contemporary culture*. Cambridge, MA: Harvard University Press.

Gallop J (1985). *Reading Lacan*. Ithaca, NY: Cornell University Press.

Hogan P C & Pandit L (1990). *Criticism and Lacan. Essays and dialogue on language, structure, and the unconscious*. Athens/London: The University of Georgia Press.

Lacan J (1966). *Écrits* (2 vols). Paris: Editions de Seuil.

Lacan J (1977). *The four fundamental concepts of psychoanalysis*. Sheridan A (trans.). London: Hogarth Press/ Institute of Psychoanalysis.

Muller J & Richardson W (1987). *Lacan and language: a reader's guide to Écrits*. New York: International Universities Press.

Ladefoged, Peter (b. 1925)

P D Fallon, Howard University, Washington, DC, USA

Peter Ladefoged is the foremost phonetician of the second half of the 20th century. Trained by David Abercrombie, Ladefoged serves as a living link between a venerable tradition of British phoneticians known for articulatorily-based description and a newer generation of analysts using computers.

Born on September 17, 1925, in Sutton, England, Ladefoged received an M.A. from the University of Edinburgh in 1951, winning the Class Medal in Phonetics. Ladefoged was lecturer in phonetics at Edinburgh from 1953 to 1959, with responsibility for the Phonetics Laboratory and for teaching experimental phonetics. He received his Ph.D. from Edinburgh in 1959. From 1961 to 1962 he was a Field Fellow in the West African Languages Survey, from which he collected material for his book (1964). Ladefoged then launched an academic career at the University of California, Los Angeles (UCLA), where he established the world-famous UCLA Phonetics Laboratory and its well-regarded UCLA Working Papers in Phonetics. Ladefoged became full professor in 1965, acting as department chair from 1977 to 1980, and has been UCLA Research Linguist/Professor of Phonetics Emeritus since 1991.

After early interests in acoustic phonetics, perceptual theories of phonetics, and work on the respiratory muscles used in speech, Ladefoged collected phonetic data on West African languages and began the lifelong task of recording and analyzing speech from little-studied languages from all inhabited continents. As Fromkin (1985: xiv) wrote, he has recorded "speech from more people speaking more languages than any other phonetician has ever done." Much of Ladefoged's work has been supported by the National Science Foundation, and the data collected has often been from endangered languages.

Laver (1991: 2) has called Ladefoged's abiding contribution "the development of a descriptive phonetic theory of spoken language whose architecture rests firmly on a foundation of empirical, objective, and quantified observation." In his 1971 work, Ladefoged tried to empirically ground phonetic features and

refers to. Without the existence of stable *signifiés*, the unconscious is the place of a sliding chain of *signifiants*.

Another of Lacan's well-known concepts is that of the 'mirror stage.' He suggests that the human infant is inherently fragmented and incomplete, a fact which is masked by the child's identification with its mirror image as a unified 'gestalt.' This identification process results in the formation of the self (*moi*) and causes the child's alienation from itself, the split subject. Once the child learns to use a language, it enters the Symbolic order, one of three domains of experience established by Lacan (the other two domains are the Real, where no language exists, and the Imaginary, the order of illusions). This symbolic order is regarded as the 'law' that governs the whole of man's sociocultural existence.

According to Lacan, language is also essential for any psychoanalytic exchange whose aim is to help the subject arrive at *parole pleine* ('full speech') via *parole vide* ('empty speech'), a necessary detour that enables the access of the true subject (*je*).

Although Lacan's style is very complex and idiosyncratic (and defies reading and interpretation to some extent), he had great influence, particularly on philosophy, sociology, cultural studies, and feminist theory, but also on interdisciplinary linguistic approaches such as Critical Discourse Analysis.

See also: Freud, Sigmund (1856–1939); Jakobson, Roman (1896–1982); Jakobson, Roman: Theory of the Sign; Levi-Strauss, Claude (b. 1908); Saussure, Ferdinand (-Mongin) de (1857–1913); Structuralism.

Bibliography

Dylan E (1996). *An introductory dictionary of Lacanian psychoanalysis*. London: Routledge.

Felman S (1987). *Jacques Lacan and the adventure of insight. Psychoanalysis in contemporary culture*. Cambridge, MA: Harvard University Press.

Gallop J (1985). *Reading Lacan*. Ithaca, NY: Cornell University Press.

Hogan P C & Pandit L (1990). *Criticism and Lacan. Essays and dialogue on language, structure, and the unconscious*. Athens/London: The University of Georgia Press.

Lacan J (1966). *Écrits* (2 vols). Paris: Editions de Seuil.

Lacan J (1977). *The four fundamental concepts of psychoanalysis*. Sheridan A (trans.). London: Hogarth Press/ Institute of Psychoanalysis.

Muller J & Richardson W (1987). *Lacan and language: a reader's guide to Écrits*. New York: International Universities Press.

Ladefoged, Peter (b. 1925)

P D Fallon, Howard University, Washington, DC, USA

Peter Ladefoged is the foremost phonetician of the second half of the 20th century. Trained by David Abercrombie, Ladefoged serves as a living link between a venerable tradition of British phoneticians known for articulatorily-based description and a newer generation of analysts using computers.

Born on September 17, 1925, in Sutton, England, Ladefoged received an M.A. from the University of Edinburgh in 1951, winning the Class Medal in Phonetics. Ladefoged was lecturer in phonetics at Edinburgh from 1953 to 1959, with responsibility for the Phonetics Laboratory and for teaching experimental phonetics. He received his Ph.D. from Edinburgh in 1959. From 1961 to 1962 he was a Field Fellow in the West African Languages Survey, from which he collected material for his book (1964). Ladefoged then launched an academic career at the University of California, Los Angeles (UCLA), where he established the world-famous UCLA Phonetics Laboratory and its well-regarded UCLA Working Papers in Phonetics. Ladefoged became full professor in 1965, acting as department chair from 1977 to 1980, and has been UCLA Research Linguist/Professor of Phonetics Emeritus since 1991.

After early interests in acoustic phonetics, perceptual theories of phonetics, and work on the respiratory muscles used in speech, Ladefoged collected phonetic data on West African languages and began the lifelong task of recording and analyzing speech from little-studied languages from all inhabited continents. As Fromkin (1985: xiv) wrote, he has recorded "speech from more people speaking more languages than any other phonetician has ever done." Much of Ladefoged's work has been supported by the National Science Foundation, and the data collected has often been from endangered languages.

Laver (1991: 2) has called Ladefoged's abiding contribution "the development of a descriptive phonetic theory of spoken language whose architecture rests firmly on a foundation of empirical, objective, and quantified observation." In his 1971 work, Ladefoged tried to empirically ground phonetic features and

phonetic theory in known contrasts in languages all over the world. The latest statement of this goal is the 1996 publication by Ladefoged and Maddieson.

Ladefoged is also a pedagogically effective professor. He won a Distinguished Teaching Award from UCLA in 1971, and his (1975) textbook is a classic and has trained a generation of linguists. More recent textbooks (1962 [2nd edn., 1996], 2001, 2003) incorporate new insights, techniques, and computers. In addition, Ladefoged was instrumental in making publically available a collection of sound files titled 'The Sounds of the World's Languages (now called 'UCLA Phonetic Data').'

Ladefoged has served as dissertation advisor for such prominent phoneticians as Victoria Fromkin, John Ohala, Ian Maddieson, and many others. In 1978, he served as president of the Linguistic Society of America. He was the recipient of a Gold Medal awarded at the XIIth International Congress of Phonetic Sciences in 1991, honoring his contributions to phonetics. He served as president of the International Phonetic Association and is currently an editor of its journal. Ladefoged also served as a technical consultant for George Cukor's film 'My Fair Lady,' for which he illustrated Henry Sweet's phonetic alphabet and helped equip Henry Higgins' lab.

See also: International Phonetic Association; Abercrombie, David (1909–1992); Sweet, Henry (1845–1912).

Bibliography

Fromkin V A (ed.) (1985). *Phonetic linguistics: essays in honor of Peter Ladefoged.* Orlando: Academic Press.
Ladefoged P (1962, 2nd edn., 1996). *Elements of acoustic phonetics.* Chicago: University of Chicago Press.
Ladefoged P (1964). *A phonetic study of West African languages.* Cambridge: Cambridge University Press.
Ladefoged P (1967). *Three areas of experimental phonetics: an auditory–instrumental survey.* Oxford: Oxford University Press.
Ladefoged P (1971). *Preliminaries to linguistic phonetics.* Chicago: University of Chicago Press.
Ladefoged P (1975). *A course in phonetics.* New York: Harcourt Brace Jovanovich.
Ladefoged P (1991). 'Remembering David Abercrombie.' *Journal of the International Phonetic Association 22,* 66–67.
Ladefoged P (2001). *Vowels and consonants: an introduction to the sounds of languages.* Malden, MA: Blackwell.
Ladefoged P (2002). Homepage. http://www.linguistics.ucla.edu/people/ladefosge/index.html.
Ladefoged P (2003). *Phonetic data analysis: an introduction to fieldwork and instrumental techniques.* Malden, MA: Blackwell.
Ladefoged P, Glick R & Criper C (1969). *Language in Uganda.* Nairobi: Oxford University Press.
Ladefoged P & Maddieson I (1996). *The sounds of the world's languages.* Oxford: Blackwell.
Laver J (1991). 'A tribute to Peter Ladefoged.' *Journal of the International Phonetic Association 21,* 1–3.

Lado, Robert (1915–1995)

M Thomas, Boston College, Chestnut Hill, MA, USA

Robert Lado was born in Florida, but his family returned to their native Spain while he was a child. When the Spanish Civil War broke out, Lado re-immigrated to the United States at age 21. He learned English while working as a sign painter, then attended college on a scholarship. Graduating with a major in English, Lado decided on a career in teaching. In graduate school he studied with University of Michigan linguist Charles C. Fries. Lado's 1950 Ph.D. thesis was on testing the proficiency of Spanish-speaking learners of English.

Lado stayed on to work with Fries as a faculty member and administrator at Michigan's English Language Institute. The ELI trained students from abroad in English as a second language, as part of a U.S. State Department–sponsored exchange program. Fries and Lado drew on midcentury American descriptive linguistics to develop language-teaching materials and techniques they called the Oral Approach. Lado created language tests and textbooks, and wrote on testing, pedagogy, and second language acquisition. His most famous publications (Lado, 1957, 1964) introduced teachers to descriptivist procedures for analyzing a language's sound system, grammar, and vocabulary. Lado argued that as a prerequisite to effective pedagogy, a teacher needs to compare the native and target languages (and cultures).

In 1960 Lado moved to Georgetown University, first as director of an ELI-like institute, later as Dean of the School of Languages and Linguistics and founder of what became Georgetown University Press. Even after retiring from full-time teaching in 1980, Lado continued his professional activities, publishing textbooks and pedagogical manuals, establishing a commercial English-language school (Lado International College), and pursuing a new research interest in teaching reading to preschoolers. Lado's

phonetic theory in known contrasts in languages all over the world. The latest statement of this goal is the 1996 publication by Ladefoged and Maddieson.

Ladefoged is also a pedagogically effective professor. He won a Distinguished Teaching Award from UCLA in 1971, and his (1975) textbook is a classic and has trained a generation of linguists. More recent textbooks (1962 [2nd edn., 1996], 2001, 2003) incorporate new insights, techniques, and computers. In addition, Ladefoged was instrumental in making publically available a collection of sound files titled 'The Sounds of the World's Languages (now called 'UCLA Phonetic Data').'

Ladefoged has served as dissertation advisor for such prominent phoneticians as Victoria Fromkin, John Ohala, Ian Maddieson, and many others. In 1978, he served as president of the Linguistic Society of America. He was the recipient of a Gold Medal awarded at the XIIth International Congress of Phonetic Sciences in 1991, honoring his contributions to phonetics. He served as president of the International Phonetic Association and is currently an editor of its journal. Ladefoged also served as a technical consultant for George Cukor's film 'My Fair Lady,' for which he illustrated Henry Sweet's phonetic alphabet and helped equip Henry Higgins' lab.

See also: International Phonetic Association; Abercrombie, David (1909–1992); Sweet, Henry (1845–1912).

Bibliography

Fromkin V A (ed.) (1985). *Phonetic linguistics: essays in honor of Peter Ladefoged.* Orlando: Academic Press.

Ladefoged P (1962, 2nd edn., 1996). *Elements of acoustic phonetics.* Chicago: University of Chicago Press.

Ladefoged P (1964). *A phonetic study of West African languages.* Cambridge: Cambridge University Press.

Ladefoged P (1967). *Three areas of experimental phonetics: an auditory–instrumental survey.* Oxford: Oxford University Press.

Ladefoged P (1971). *Preliminaries to linguistic phonetics.* Chicago: University of Chicago Press.

Ladefoged P (1975). *A course in phonetics.* New York: Harcourt Brace Jovanovich.

Ladefoged P (1991). 'Remembering David Abercrombie.' *Journal of the International Phonetic Association 22,* 66–67.

Ladefoged P (2001). *Vowels and consonants: an introduction to the sounds of languages.* Malden, MA: Blackwell.

Ladefoged P (2002). Homepage. http://www.linguistics.ucla.edu/people/ladefosge/index.html.

Ladefoged P (2003). *Phonetic data analysis: an introduction to fieldwork and instrumental techniques.* Malden, MA: Blackwell.

Ladefoged P, Glick R & Criper C (1969). *Language in Uganda.* Nairobi: Oxford University Press.

Ladefoged P & Maddieson I (1996). *The sounds of the world's languages.* Oxford: Blackwell.

Laver J (1991). 'A tribute to Peter Ladefoged.' *Journal of the International Phonetic Association 21,* 1–3.

Lado, Robert (1915–1995)

M Thomas, Boston College, Chestnut Hill, MA, USA

Robert Lado was born in Florida, but his family returned to their native Spain while he was a child. When the Spanish Civil War broke out, Lado re-immigrated to the United States at age 21. He learned English while working as a sign painter, then attended college on a scholarship. Graduating with a major in English, Lado decided on a career in teaching. In graduate school he studied with University of Michigan linguist Charles C. Fries. Lado's 1950 Ph.D. thesis was on testing the proficiency of Spanish-speaking learners of English.

Lado stayed on to work with Fries as a faculty member and administrator at Michigan's English Language Institute. The ELI trained students from abroad in English as a second language, as part of a U.S. State Department–sponsored exchange program. Fries and Lado drew on midcentury American descriptive linguistics to develop language-teaching materials and techniques they called the Oral Approach. Lado created language tests and textbooks, and wrote on testing, pedagogy, and second language acquisition. His most famous publications (Lado, 1957, 1964) introduced teachers to descriptivist procedures for analyzing a language's sound system, grammar, and vocabulary. Lado argued that as a prerequisite to effective pedagogy, a teacher needs to compare the native and target languages (and cultures).

In 1960 Lado moved to Georgetown University, first as director of an ELI-like institute, later as Dean of the School of Languages and Linguistics and founder of what became Georgetown University Press. Even after retiring from full-time teaching in 1980, Lado continued his professional activities, publishing textbooks and pedagogical manuals, establishing a commercial English-language school (Lado International College), and pursuing a new research interest in teaching reading to preschoolers. Lado's

publications (Jankowsky, 1985) include over 150 books, articles, and reviews. Rounding out a busy, creative, life, Lado and his wife Lucia raised five daughters and five sons.

With the eclipse of descriptive linguistics and its concept of second language learning, Lado is now best remembered as the author of *Linguistics across cultures*. In particular, he is represented as the architect of the 'Contrastive Analysis Hypothesis' – essentially the claim that wherever a learner's native and target languages differ, the learner will face difficulty and delay in acquiring the target language. Contrastive Analysis is typically presented (e.g., by Gass and Selinker, 2001) as a product of American descriptivism and Bloomfieldian behaviorism. Thomas (submitted) argues that this misconstrues the relationships of descriptivism, behaviorism, and Contrastive Analysis in Lado's (and Fries's) work: first, because Lado was not a behaviorist, and second, because his commitment to cross-linguistic comparison is not equivalent in extent or content to the Contrastive Analysis Hypothesis.

See also: Applied Linguistics: Overview and History; Fries, Charles Carpenter (1897–1967); Interlanguage; Language Teaching: History; Structuralism.

Bibliography

Gass S M & Selinker L (2001). *Second language acquisition: An introductory course.* (2nd edn.). Mahwah, NJ: Lawrence Erlbaum.

Jankowsky K R (ed.) (1985). *Scientific and humanistic dimensions of language: A festschrift for Robert Lado.* Amsterdam: John Benjamins.

Lado R (1957). *Linguistics across cultures.* Ann Arbor: University of Michigan.

Lado R (1964). *Language teaching: A scientific approach.* New York: McGraw-Hill.

Thomas M (submitted). 'The 'Ebeling Principle' in the historiography of generative second language acquisition.'

Lahnda

C Shackle, University of London, London, UK

The name 'Lahnda' (Panjabi *lahiṁdā* 'western'), like the more natural feminine 'Lahndi,' is an invented blanket term without local currency used to distinguish the Indo-Aryan dialects spoken in western Panjab and the adjacent regions of Pakistan (by some 30 million speakers) from the Panjabi proper native to the central and eastern districts of the Panjab. The boundary between Lahnda and Panjabi (Shackle, 2003) is anyway an uncertain one. Many common features, e.g., retained historical geminates or the possessive marker *dā*, have been reinforced by overlapping literary traditions. Since the 1990s, however, local politicocultural movements have emphasized the distinctive character of some varieties of Lahnda, and the southern Seraiki (Ser) (Saraiki) and to a lesser extent the northern Hindko (also termed Pahari or Panjistani) are beginning to be employed as independent literary languages consciously rivaling Pakistani Panjabi (Pj). Written in the Perso-Urdu script, all three incipient standards are considerably influenced by Urdu.

Siraiki

The rather homogeneous southwestern Seraiki dialects (Shackle, 1976) are notably distinguished from Panjabi by the retention of historical aspiration and the development of four implosive phonemes /ɠ, ʄ, ᶑ, ɓ/, shared with Sindhi, thus Sanskrit *baddha-* 'bound' > Pj /bə´dda/ but Ser /' ɓəddha/, Sindhi /' ɓədho/.

Distinctive Seraiki morphological features include the passive extension /-'i-/ and the sigmatic future, with stressed extension of transitive stems, e.g., Ser /kə'resi/ 'will do', passive /kə'risi/ versus Pj /kə´rega, 'kita ja´vega/; and a full set of suffixed pronouns, often entailing shifted stress, e.g., /'kita/ 'did', /'kitum/ 'I did', /kɪ'tose/ 'we did'. Common lexical distinctions include Ser /vəɲɲaɳ/ 'to go', future /væsi/, versus Pj /jaɳa, ja´vega/, and the Ser objective and ablative postpositions /kũ, kənũ/ versus Pj /nũ, tõ/. Typical Seraiki shibboleths appear in /ə'khɪomɪs jo sakũ ɓəhũ jəldi vəɲɳa posi/ [*told-me-him that us- to very quick going fall-will*] 'I told him that we should have to go very early', where /ə'khɪomɪs/ has a double suffix (/'akhɪa/ + 1SG /-m/ + 3SG /-s/), versus the analytic Pj /mæ̃ o´nũ akhɪa pəi sanũ bɔ´t jəldi jaɳa pə´vega/.

Hindko

Hindko is the term locally used to cover the heterogeneous northern dialects spoken in the hilly areas above the Salt Range (Shackle, 1980). These include the well-described Awankari (Aw) (Bahri, 1963) and the very different Hindko of Peshawar city (Pe). Like

publications (Jankowsky, 1985) include over 150 books, articles, and reviews. Rounding out a busy, creative, life, Lado and his wife Lucia raised five daughters and five sons.

With the eclipse of descriptive linguistics and its concept of second language learning, Lado is now best remembered as the author of *Linguistics across cultures*. In particular, he is represented as the architect of the 'Contrastive Analysis Hypothesis' – essentially the claim that wherever a learner's native and target languages differ, the learner will face difficulty and delay in acquiring the target language. Contrastive Analysis is typically presented (e.g., by Gass and Selinker, 2001) as a product of American descriptivism and Bloomfieldian behaviorism. Thomas (submitted) argues that this misconstrues the relationships of descriptivism, behaviorism, and Contrastive Analysis in Lado's (and Fries's) work: first, because Lado was not a behaviorist, and second, because his commitment to cross-linguistic comparison is not equivalent in extent or content to the Contrastive Analysis Hypothesis.

See also: Applied Linguistics: Overview and History; Fries, Charles Carpenter (1897–1967); Interlanguage; Language Teaching: History; Structuralism.

Bibliography

Gass S M & Selinker L (2001). *Second language acquisition: An introductory course.* (2nd edn.). Mahwah, NJ: Lawrence Erlbaum.

Jankowsky K R (ed.) (1985). *Scientific and humanistic dimensions of language: A festschrift for Robert Lado.* Amsterdam: John Benjamins.

Lado R (1957). *Linguistics across cultures.* Ann Arbor: University of Michigan.

Lado R (1964). *Language teaching: A scientific approach.* New York: McGraw-Hill.

Thomas M (submitted). 'The 'Ebeling Principle' in the historiography of generative second language acquisition.'

Lahnda

C Shackle, University of London, London, UK

The name 'Lahnda' (Panjabi *lahiṁdā* 'western'), like the more natural feminine 'Lahndi,' is an invented blanket term without local currency used to distinguish the Indo-Aryan dialects spoken in western Panjab and the adjacent regions of Pakistan (by some 30 million speakers) from the Panjabi proper native to the central and eastern districts of the Panjab. The boundary between Lahnda and Panjabi (Shackle, 2003) is anyway an uncertain one. Many common features, e.g., retained historical geminates or the possessive marker *dā*, have been reinforced by overlapping literary traditions. Since the 1990s, however, local politicocultural movements have emphasized the distinctive character of some varieties of Lahnda, and the southern Seraiki (Ser) (Saraiki) and to a lesser extent the northern Hindko (also termed Pahari or Panjistani) are beginning to be employed as independent literary languages consciously rivaling Pakistani Panjabi (Pj). Written in the Perso-Urdu script, all three incipient standards are considerably influenced by Urdu.

Siraiki

The rather homogeneous southwestern Seraiki dialects (Shackle, 1976) are notably distinguished from Panjabi by the retention of historical aspiration and the development of four implosive phonemes /ɠ, ʄ, ᶑ, ɓ/, shared with Sindhi, thus Sanskrit *baddha-* 'bound' > Pj /bə´dda/ but Ser /' ɓəddha/, Sindhi /' ɓədho/.

Distinctive Seraiki morphological features include the passive extension /-'i-/ and the sigmatic future, with stressed extension of transitive stems, e.g., Ser /kə'resi/ 'will do', passive /kə'risi/ versus Pj /kə´rega, 'kita ja´vega/; and a full set of suffixed pronouns, often entailing shifted stress, e.g., /'kita/ 'did', /'kitum/ 'I did', /kɪ'tose/ 'we did'. Common lexical distinctions include Ser /vəɲɲaɳ/ 'to go', future /væsi/, versus Pj /jaɳa, ja´vega/, and the Ser objective and ablative postpositions /kũ, kənũ/ versus Pj /nũ, tõ/. Typical Seraiki shibboleths appear in /ə'khɪoᴍɪs jo sakũ ɓəhũ jəldi vəɲɳa posi/ [*told-me-him that us- to very quick going fall-will*] 'I told him that we should have to go very early', where /ə'khɪoᴍɪs/ has a double suffix (/'akhɪa/ + 1SG /-m/ + 3SG /-s/), versus the analytic Pj /mæ̃ o´nũ akhɪa pəi sanũ bɔ´t jəldi jaɳa pə´vega/.

Hindko

Hindko is the term locally used to cover the heterogeneous northern dialects spoken in the hilly areas above the Salt Range (Shackle, 1980). These include the well-described Awankari (Aw) (Bahri, 1963) and the very different Hindko of Peshawar city (Pe). Like

Panjabi, Hindko has tonal realizations of historical aspiration, but the phonetic features associated with the Panjabi low-rising tone accompany a high-falling tone in Peshawar Hindko, e.g., Sanskrit *bhāra-* 'load' → Aw, Ser /bhar/, versus Pe /p'àr/ (Pj /p'a΄r/). The Seraiki sigmatic future and pronominal suffixes are shared by Hindko, where the model sentence would appear as Aw /mæ̃ usã akha jo bə΄ũ jəldi vəɲɳa posi/, Pe /məne unũ ki΄a ke bɔ΄t jəldi jaɳa pæsi/.

See also: Pakistan: Language Situation; Punjabi.

Bibliography

Bahri H (1963). *Lahndi phonetics.* Allahabad, India: Bharati Press.

Shackle C (1976). *The Siraiki language of central Pakistan: a reference grammar.* London: SOAS.

Shackle C (1980). 'Hindko in Kohat and Peshawar.' *Bulletin of SOAS 43,* 482–510.

Shackle C (2003). 'Panjabi.' In Cardona G (ed.) *The Indo-Aryan languages.* London: Routledge.

Lak

V A Friedman, University of Chicago, Chicago, IL, USA

Introduction

The Lak language (*Ethnologue* code LBE) belongs to the Daghestanian branch of the Nakh-Daghestanian family and has over 200 000 speakers, mostly in the Republic of Daghestan, Russian Federation (maps of the region are available through the website 'Thesaurus Indogermanischer Text- und Sprachmaterialien (TITUS)'; see Relevant Websites, at the end of this article). The Lak self-designation is *Lak* (adjective *Lak:u*); other terms include Turkish *Beyaz lezgi* 'White Lezgian' and *Kazikumux*, after Kumux, the main aul (village) of Lakkia and the former center of a feudal state. Lak glosses started around 1600; whole texts appeared in 1734. Lak was written in the Arabic Alphabet until 1928, then in Latin until 1938, and finally in Cyrillic. Lak has five dialects; the standard language is based on Kumux.

Phonology

The Lak vowels are /a, i, u/, all of which can be distinctively pharyngealized, which results in their allophonic centralization [æˤ, eˤ, øˤ]. The Lak consonants are shown in **Table 1**. Geminate (emphatic) consonants are realized as simple unaspirated, except in prevocalic (and, for stops, noninitial) position. Consonant labialization is distinctive in some dialects, and vowel length and stress interact.

Table 1 Lak consonants

b	p	p:	p'							w		m
d	t	t:	t'	c	c:	c'	z	s	s:	r	l	n
				č	č:	č'	ž	š	š:	j		
g	k	k:	k'					x̂	x̂:			
	q	q:	q'				ğ	x̌	x̌:			
								ħ				
	ʔ							h				

Morphology

Lak has four noun classes: I = male sentient, II = mature female sentient, III = other animate and some inanimate, and IV = inanimate and a few lower animals. Any part of speech can take class agreement markers, which are prefixed, infixed, and/or suffixed. Lak nouns have four stems, e.g., nominative (= absolutive) singular, as in *q:at:a* 'house'; oblique singular, as in *q:at-lu-*; nominative plural, as in *q:at-ri*; and oblique plural, as in *q:at-ra-*. There are more than 30 stem formants. Case endings attach agglutinatively to the oblique. The three core cases are nominative-Ø, genitive *-l*, and dative *-n*. Secondary cases include addressive/possessive *-x̂*, admotive/dative -2-*x̌:un*, ablative/involuntative/possibilitive *-š:a*, and comitative *-š:al*. A few other affixes attach to the oblique and/or nominative stems.

Lak has six oblique locational affixes (LA) to which five movement affixes (MA) can attach, forming potentially 30 local cases. Not all local cases occur in all combinations or with all substantives, but, unlike spatial postpositions (which take the genitive), local cases have nonlocational functions and cannot govern across a conjunction (abbreviations: IMP, imperative; SG, singular; NOM, nominative; GEN, genitive; EMPH, emphatic; OBL, oblique; PL, plural; INF, infinitive).

oˤwč-a   wi-l-a   dus-tura-j-n
invite-IMP   you.SG-GEN-EMPH   friend-OBL.PL-on.LA-to.MA
wa   malla-na-j-n   pulaw   b-uk-an
and   mullah-OBL-on.LA-to.MA   pilaf(III)   III-eat-INF
'Invite your friends and the mullah to eat pilaf'.

Panjabi, Hindko has tonal realizations of historical aspiration, but the phonetic features associated with the Panjabi low-rising tone accompany a high-falling tone in Peshawar Hindko, e.g., Sanskrit *bhāra-* 'load' → Aw, Ser /bhar/, versus Pe /p'àr/ (Pj /p'a′r/). The Seraiki sigmatic future and pronominal suffixes are shared by Hindko, where the model sentence would appear as Aw /mæ̃ usã akha jo bə′ũ jəldi vəɲɳa posi/, Pe /məne unũ ki′a ke bɔ′t jəldi jaɳa pæsi/.

See also: Pakistan: Language Situation; Punjabi.

Bibliography

Bahri H (1963). *Lahndi phonetics.* Allahabad, India: Bharati Press.

Shackle C (1976). *The Siraiki language of central Pakistan: a reference grammar.* London: SOAS.

Shackle C (1980). 'Hindko in Kohat and Peshawar.' *Bulletin of SOAS 43,* 482–510.

Shackle C (2003). 'Panjabi.' In Cardona G (ed.) *The Indo-Aryan languages.* London: Routledge.

Lak

V A Friedman, University of Chicago, Chicago, IL, USA

Introduction

The Lak language (*Ethnologue* code LBE) belongs to the Daghestanian branch of the Nakh-Daghestanian family and has over 200 000 speakers, mostly in the Republic of Daghestan, Russian Federation (maps of the region are available through the website 'Thesaurus Indogermanischer Text- und Sprachmaterialien (TITUS)'; see Relevant Websites, at the end of this article). The Lak self-designation is *Lak* (adjective *Lak:u*); other terms include Turkish *Beyaz lezgi* 'White Lezgian' and *Kazikumux*, after Kumux, the main aul (village) of Lakkia and the former center of a feudal state. Lak glosses started around 1600; whole texts appeared in 1734. Lak was written in the Arabic Alphabet until 1928, then in Latin until 1938, and finally in Cyrillic. Lak has five dialects; the standard language is based on Kumux.

Phonology

The Lak vowels are /a, i, u/, all of which can be distinctively pharyngealized, which results in their allophonic centralization [æˁ, eˁ, øˁ]. The Lak consonants are shown in Table 1. Geminate (emphatic) consonants are realized as simple unaspirated, except in prevocalic (and, for stops, noninitial) position. Consonant labialization is distinctive in some dialects, and vowel length and stress interact.

Table 1 Lak consonants

b	p	p:	p'							w		m
d	t	t:	t'	c	c:	c'	z	s	s:	r	l	n
				č	č:	č'	ž	š	š:	j		
g	k	k:	k'					x̂	x̂:			
	q	q:	q'				ğ	x̌	x̌:			
								ħ				
	ʔ							h				

Morphology

Lak has four noun classes: I = male sentient, II = mature female sentient, III = other animate and some inanimate, and IV = inanimate and a few lower animals. Any part of speech can take class agreement markers, which are prefixed, infixed, and/or suffixed. Lak nouns have four stems, e.g., nominative (= absolutive) singular, as in *q:at:a* 'house'; oblique singular, as in *q:at-lu-*; nominative plural, as in *q:at-ri*; and oblique plural, as in *q:at-ra-*. There are more than 30 stem formants. Case endings attach agglutinatively to the oblique. The three core cases are nominative-Ø, genitive *-l*, and dative *-n*. Secondary cases include addressive/possessive *-x̂*, admotive/dative -2-*x̌:un*, ablative/involuntative/possibilitive *-š:a*, and comitative *-š:al*. A few other affixes attach to the oblique and/or nominative stems.

Lak has six oblique locational affixes (LA) to which five movement affixes (MA) can attach, forming potentially 30 local cases. Not all local cases occur in all combinations or with all substantives, but, unlike spatial postpositions (which take the genitive), local cases have nonlocational functions and cannot govern across a conjunction (abbreviations: IMP, imperative; SG, singular; NOM, nominative; GEN, genitive; EMPH, emphatic; OBL, oblique; PL, plural; INF, infinitive).

oˁwč-a wi-l-a dus-tura-j-n
invite-IMP you.SG-GEN-EMPH friend-OBL.PL-on.LA-to.MA
wa malla-na-j-n pulaw b-uk-an
and mullah-OBL-on.LA-to.MA pilaf(III) III-eat-INF
'Invite your friends and the mullah to eat pilaf'.

Lak has five sets of deictics, which also serve as third-person markers: *wa*, near the speaker, new information; *mu*, relevant to the addressee, old information; *ta*, opposite, level (older unmarked); *ga*, unmarked (older below speaker); and *k'a*, above speaker. The verb has three aspectual stems, perfective/unmarked (*buc-in* 'bring-INF'), durative in *-la-* (*buc-la-n*), and iterative in *-awa-* (*buc-awa-n*). Synthetic forms of the marked aspects occur only in the present and future. The infinitive serves as the base for the future. Past tense forms usually have a class marker infixed before the last root consonant, and an infixed negator (indicative *q:a-*, imperative *ma-*) precedes the infixed class marker. The verb has numerous synthetic and analytic paradigms marking aspect, tense, mood, and evidential, some with marking for person as well as class and number.

Syntax

Lak is basically object-verb, attributive/genitive-head; it has pragmatically conditioned free word order and a mixed ergative/accusative system. The converbs *ban* 'to do' and *x̂un* 'to become' are the most frequent markers of transitive/causative vs. intransitive, respectively. For the agent of an ordinary transitive verb, personal pronouns (first and second person) remain in the nominative; other agents take the genitive, which also functions as ergative. Case assignment and verb agreement also depend on the semantics of the verb, focus, and the pragmatic implications of the clause. Experiencers take dative; ability and accident are marked by ablative. Complement clauses trigger class III agreement. The following sentences are illustrative (a resumed morpheme interrupted by a class marker is indicated by <$>; GER, gerund; PRES, present; ABL, ablative; DAT, dative; ABS, absolutive; DUR, durative aspect; PA, perfective aspect).

ga-na-l k'ili d-a-r-x̌:-unu
he-OBL(I)-GEN saddle(IV) IV-sell-IV-$-PAST.GER
Ø-u-r
I-be-3SG.PRES
'He has sold the saddle'.

ga-na-l k'ili d-a-r-x̌:-unu
he-OBL(I)-GEN saddle(IV) IV-sell-IV-$-PAST.GER
d-u-r
IV-be-3SG.PRES
'Apparently he sold the saddle'.

ga-na-š:a k'ili d-a-r-x̌:-unu
he-OBL(I)-ABL saddle(IV) IV-sell-IV-$-PAST.GER
d-u-r
IV-be-3SG.PRES
'He accidentally sold the saddle'.

ga-na-n k'ili d-ax̌:an č:a-j
he-OBL(I)-DAT saddle(IV) IV-sell-INF want-PRES.GER
b-u-r
III-be-3SG.PRES
'He wants to sell the saddle'.

ga-na-š·a k'ili d-ax̌:-an
he-OBL(I)-ABL saddle (IV) IV-sell-INF
b-uˁq-la-j b-u-r
III-can-DUR-PRES.GER III-be-3SG.PRES
'He can sell the saddle'.

ni-t:i-l q:at-lu-w-un-m-aj
mother(II)-OBL-GEN house-OBL-in(LA)-toward(MA)-III-$
č:at' la-w-s-un na-j
bread(III) bring-III-$-PA.GER go-PRES.GER
b-u-r
III-be-3SG.PRES
'Mother brings bread into the house'.

ninu q:at-lu-wun-m-aj
mother(II).ABS house-OBL-in(LA)-toward(MA)-III-$
č:at' la-w-sun na-j
bread(III) bring-III-$-PA.GER go-PRES.GER
d-u-r
II-be-3SG.PRES
'Bread is brought by mother into the house'.

ninu q:at-lu-wun-n-aj
mother(II).ABS house-OBL-in(LA)-toward(MA)-II-$
č:at' la-w-sun na-j d-u-r
bread(III) bring-III-$-PA.GER go-PRES.GER II-be-3SG.PRES
'It is mother who brings bread into the house'.

See also: Caucasian Languages; Russian Federation: Language Situation.

Bibliography

Anderson G (1997). 'Lak phonology.' In Kaye A (ed.) *Phonologies of Asia and Africa* (including the Caucasus). Winona Lake, IN: Eisenbrauns. 973–997.

Burč'uladze G (1986). *Brunvata Šedgenilobisa da arsebit saxelta brunebis p'rocesebis dziritadi sak'itxebi lak'ur enaši* [*Fundamental questions of case composition and processes of declension of the noun in the Lak language*]. Tbilisi: Mecniereba.

Burč'uladze G (1987). *Lak'uri zmna* [The Lak verb]. Tbilisi: Mecniereba.

Friedman V (1992). 'Lak substantival declension: 40 cases or 50?' In Aronson H (ed.) *The non-Slavic languages of the USSR: linguistic studies – second series.* Chicago: Chicago Linguistic Society. 113–133.

Friedman V (1994). '*Ga* in Lak and the three "there"s: deixis and markedness in Daghestan.' In Aronson H (ed.) *NSL.7: linguistic studies in the non-Slavic languages of the commonwealth of independent states and the Baltic republics.* Chicago: Chicago Linguistic Society. 79–93.

Friedman V (2004). 'The expression of speaker subjectivity in Lak.' In Guentchéva Z (ed.) *L'Énonciation médiatisée 2*. Louvain: Peeters. 75–83.
Helmbrecht J (1996). 'The syntax of personal agreement in East-Caucasian languages.' *Sprachtypologie und Universalienforschung 49*, 127–148.
Kazenin K (1998). 'On patient demotion in Lak.' In Kulikov L & Vater H (eds.) *Typology of verbal categories*. Tübingen: Max Niemeyer. 95–115.
Kazenin K (2001). 'Focus in Daghestanian languages and word order parameter.' *Linguistic Typology 1*, 289–316.
Khajdakov S (1966). 'The dialect-divisions of Lak.' *Studia Caucasica 2*, 9–18.
Khalilov Kh (1976). *Jumor I satira narodov Dagestana* [*Humor and satire of the peoples of Daghestan*]. Maxačkala: AN SSSR.
Kibrik A E (1997). 'Beyond subject and object: toward a comprehensive relational typology.' *Journal of Linguistic Typology 1*, 279–346.
Murkelinskij G (1971). *Grammatika lakskogo jazyka.* [Grammar of the Lak language] Maxačkala: Dagučpedgiz.
Žirkov L (1955). *Lakskij jazyk* [The Lak language]. Moscow: AN SSSR.

Relevant Websites

http://titus.fkidg1.uni-frankfurt.de – The website of Thesaurus Indogermanischer Text- und Sprachmaterialien (TITUS). Maps of Asia and the Caucasus region are available by selecting the links 'Teaching Materials,' then 'Language Maps'.

Lakota

B Ingham, SOAS, London, UK

Introduction

Lakota is one of a group of closely related dialects sometimes referred to by linguists as Dakotan. These include Lakota in the west, Dakota in the east, Nakota in the north, and Nakoda in the northwest. The speakers of Lakota and Dakota were traditionally referred to in English as the Sioux, those of Nakota as the Assiniboine, and those of Nakoda as the Stoney. Lakota and Dakota are mutually intelligible. Reports differ as to how far Nakota and Nakoda are intelligible with the other two. Dakotan is part of a group of languages known as Siouan-Caddoan centered mainly on the central plains and prairies, but also represented on the eastern seaboard.

Reports on the number of speakers of Lakota range from 6000 to 20 000. Great efforts are being made to preserve the language in schools, colleges, and universities in the region and there is probably a considerable degree of partial or receptive knowledge of it.

Morphology

Major word classes of Lakota are noun, verb, adverb, postposition, demonstrative, pronoun, and conjunction. The verb in particular can be regarded as polysynthetic and noun incorporation occurs in the verb and adverb. The functions often covered by adjectives in other languages are covered in Lakota by stative verbs and adverbs.

The verb system is of the split intransitive type, where agents occur only in the Active verb class while patients occur with the Active and Stative types. These are marked in the verb by prefixes or infixes as shown below:

agent marker	patient marker
wa-'I'	ma-'me'
ya-'you'	ni-'you'
uŋ(k)-'we' ('you and I')	uŋ(k)-'us' ('you and me')
	wicʻa-'them' (animate)

composite pronoun prefix
cʻi-'I (agent)-you (patient)'

Plurality in the third and second person and inclusiveness of third persons in the first plural is marked by a suffix *-pi*. The occurrence of these markers with the two verb types is shown below:

Active verb *u* 'to come'

sing, exclusive		plur, inclusive	
wau	'I come'		
yau	'you (sing) come'	yaupi	'you (plur) come'
u	'he, she, it comes'	upi	'they (animate) come'
uŋku	'we (excl) come'	uŋkupi	'we (incl) come'

Stative verb *kʻuja* 'be ill'

sing, exclusive		plur, inclusive	
makʻuje	'I am ill'		
nikʻuje	'you (sing) are ill'	nikʻujapi	'you (plur) are ill'
kʻuje	'he is ill'	kʻujapi	'they are ill'
uŋkʻuje	'we (excl) are ill'	uŋkʻujapi	'we (incl) are ill'

Friedman V (2004). 'The expression of speaker subjectivity in Lak.' In Guentchéva Z (ed.) *L'Énonciation médiatisée 2*. Louvain: Peeters. 75–83.
Helmbrecht J (1996). 'The syntax of personal agreement in East-Caucasian languages.' *Sprachtypologie und Universalienforschung 49*, 127–148.
Kazenin K (1998). 'On patient demotion in Lak.' In Kulikov L & Vater H (eds.) *Typology of verbal categories*. Tübingen: Max Niemeyer. 95–115.
Kazenin K (2001). 'Focus in Daghestanian languages and word order parameter.' *Linguistic Typology 1,* 289–316.
Khajdakov S (1966). 'The dialect-divisions of Lak.' *Studia Caucasica 2,* 9–18.
Khalilov Kh (1976). *Jumor I satira narodov Dagestana* [*Humor and satire of the peoples of Daghestan*]. Maxačkala: AN SSSR.
Kibrik A E (1997). 'Beyond subject and object: toward a comprehensive relational typology.' *Journal of Linguistic Typology 1,* 279–346.
Murkelinskij G (1971). *Grammatika lakskogo jazyka.* [Grammar of the Lak language] Maxačkala: Dagučpedgiz.
Žirkov L (1955). *Lakskij jazyk* [The Lak language]. Moscow: AN SSSR.

Relevant Websites

http://titus.fkidg1.uni-frankfurt.de – The website of Thesaurus Indogermanischer Text- und Sprachmaterialien (TITUS). Maps of Asia and the Caucasus region are available by selecting the links 'Teaching Materials,' then 'Language Maps'.

Lakota

B Ingham, SOAS, London, UK

Introduction

Lakota is one of a group of closely related dialects sometimes referred to by linguists as Dakotan. These include Lakota in the west, Dakota in the east, Nakota in the north, and Nakoda in the northwest. The speakers of Lakota and Dakota were traditionally referred to in English as the Sioux, those of Nakota as the Assiniboine, and those of Nakoda as the Stoney. Lakota and Dakota are mutually intelligible. Reports differ as to how far Nakota and Nakoda are intelligible with the other two. Dakotan is part of a group of languages known as Siouan-Caddoan centered mainly on the central plains and prairies, but also represented on the eastern seaboard.

Reports on the number of speakers of Lakota range from 6000 to 20 000. Great efforts are being made to preserve the language in schools, colleges, and universities in the region and there is probably a considerable degree of partial or receptive knowledge of it.

Morphology

Major word classes of Lakota are noun, verb, adverb, postposition, demonstrative, pronoun, and conjunction. The verb in particular can be regarded as polysynthetic and noun incorporation occurs in the verb and adverb. The functions often covered by adjectives in other languages are covered in Lakota by stative verbs and adverbs.

The verb system is of the split intransitive type, where agents occur only in the Active verb class while patients occur with the Active and Stative types. These are marked in the verb by prefixes or infixes as shown below:

agent marker	patient marker
wa-'I'	ma-'me'
ya-'you'	ni-'you'
uŋ(k)-'we' ('you and I')	uŋ(k)-'us' ('you and me')
	wicʿa-'them' (animate)

composite pronoun prefix
cʿi-'I (agent)-you (patient)'

Plurality in the third and second person and inclusiveness of third persons in the first plural is marked by a suffix *-pi*. The occurrence of these markers with the two verb types is shown below:

Active verb *u* 'to come'

sing, exclusive		plur, inclusive	
wau	'I come'		
yau	'you (sing) come'	yaupi	'you (plur) come'
u	'he, she, it comes'	upi	'they (animate) come'
uŋku	'we (excl) come'	uŋkupi	'we (incl) come'

Stative verb *kʿuja* 'be ill'

sing, exclusive		plur, inclusive	
makʿuje	'I am ill'		
nikʿuje	'you (sing) are ill'	nikʿujapi	'you (plur) are ill'
kʿuje	'he is ill'	kʿujapi	'they are ill'
uŋkʿuje	'we (excl) are ill'	uŋkʿujapi	'we (incl) are ill'

Forms with both agent and patient prefixes, showing some infixed, are shown with the active verb *ole* ‘to seek’: *owicʽawale* ‘I seek them’, *owicʽayale* ‘you (sing) seek them’, *ocʽile* ‘I seek you (sing)’, *ocʽilepi* ‘I seek you (plur)’, *onile* ‘he seeks you (sing)’, *uŋkonilepi* ‘we (incl) seek you’, *omayale* ‘you (sing) seek me’, *uŋkoyalepi* ‘you (plur) seek us (incl)’.

Nouns do not show number distinctions, but may mark a possessor as in *mi-hiŋgna* ‘my husband’, *ni-hiŋgna* ‘your husband’, *mi-cʽiŋkṡi* ‘my son’, *cʽiŋkṡi-tku* ‘his, her son’. Other nominal prefixes may distinguish noun classes such as instruments as in *wi-cʽapʽe* ‘fork’, abstract concepts as in *wo-slolye* ‘knowledge’, time and space concepts as in *o-apʽe* ‘hour’, *o-makʽa* ‘year’, *o-napʽe* ‘refuge’.

Word Derivation

Word derivation is by affixation and compounding. Adverbs are formed with a suffix *-ya*, *-yela* or *-yaŋ* from stative verbs as in *waṡte* ‘be good’, *waṡte-ya* ‘well’, *ska* ‘be white’, *ska-yela* ‘whitely’, *wakʽaŋ* ‘be sacred’, *wakʽaŋ-yaŋ* ‘in a sacred manner’. Adverbs are widely used to state characteristics of objects often with the verb *haŋ/he* ‘stand’ as in *maḣpiya waŋ ska-yela he* ‘a cloud stood whitely’ meaning ‘there was a white cloud’, *op’osya he* ‘it stood coldly and clearly’ meaning ‘the weather was cold and clear’. Postpositions can be formed from adverbs by a prefix *i-* as in *wiyoḣpeyata* ‘in the west’, *i-wiyoḣ-peyata* ‘to the west of’.

A set of circumstantial stems are important in forming verbs, adverbs, wh-words and postpositions as in *tokʽeca* ‘be somehow’, *hecʽeca* ‘be like this’, *iyecʽeca* ‘be like’, *tokʽel* ‘how’, *hecʽel* ‘like this’, *iyecʽel* ‘like, as’ from a stem **-kʽecʽa/-kʽel* indicating ‘quality’. Other such stems are **-haŋ* ‘time’, **-kʽetu* ‘occurrence’, **-nakeca* ‘number’, **-haŋkeca* ‘(extent’, yielding among others *hehaŋ* ‘then’, *tohaŋ* ‘when’, *iyehaŋtu* ‘be time for’, *hecʽetu* ‘happen thus’, *iyecʽetu* ‘happen as’, *iyenakeca* ‘be as many as’, *iyena* ‘as many as’, *tonakeca* ‘how many’, *tohaŋyaŋ* ‘how far’, *iyehaŋyaŋ* ‘as far as’.

Syntax

Word order is generally agent-patient-verb as in *wicʽaṡa ki cʽiŋkṡitku ki ole* ‘the man looked for his son.’ Relative clauses are marked by the use of *waŋ* ‘one’, *ki* ‘the’ and *he* ‘that’ as in *wicʽaṡa waŋ cʽiŋkṡitku ki ole ki he owale* ‘I looked for the man who was looking for his son.’ Sentences can be embedded in higher sentences by using *ki/k’uŋ* ‘the’ and by certain postpositions as in *cʽiŋkṡitku ki ole ki slolwaye* ‘I know that he was looking for his son’, *cʽiŋkṡitku tʽi ekta wau* ‘I came to where his son lived’.

Nouns can be incorporated into verbs and verbs subordinated to other verbs by preposing, often also with stem truncation, as in *ṡuŋg-ole* ‘he looked for horses’, *ṡuŋg-manuŋ* ‘he stole horses’, *ṡuŋg-ole-mani* ‘he traveled looking for horses’ (*ṡuŋg-* ≪ *ṡuŋka* ‘horse, dog’), *iŋyaŋg-mani* ‘he traveled running’ (*iŋyaŋg-* ≪ *iŋyaŋka* ‘run’), *kaḣ-ṡi* ‘ask to make’ (*kaḣ-* ≪ *kaġa* ‘make’).

Men’s and Women’s Speech

These are distinguished by certain sentence-final particles of high frequency of occurrence shown below.

		male speaker	female speaker
declarative		-yelo, -welo, -kṡto	-ye, -we, -kṡto
interrogative		-hwo (*-he* informally)	-he
imperative	sing	-yo, -wo	-ye, -we
	plur	-po	-pe

See also: Caddoan Languages; Polysynthetic Language: Central Siberian Yupik; Siouan Languages; United States of America: Language Situation.

Bibliography

Boas F & Deloria E C (1941). *Dakota grammar.* Washington: U. S. Government Printing Office.

Buechel E (1939). *A grammar of Lakota: the language of the Teton Sioux Indians.* Rosebud SD: Rosebud Educational Society.

Buechel E (1970). *Lakota English dictionary.* Pine Ridge SD: Red Cloud Indian School.

Buechel E (1978). *Lakota tales and texts.* Pine Ridge SD: Red Cloud Indian School.

Bushotter G (1887). *Bushotter papers: texts by George Bushotter 1887.* Manuscript 4800 Dorsey Papers: Dakota: Teton (3.1.1.3) 103. Smithsonian Institution National Anthropological Archives.

Deloria E C (1932). *Dakota texts.* Publications of the American Ethnology Society. Boas F (ed.). New York: G. E. Stechert. Reprinted in 1974. New York: AMS Press.

Ingham B (2001). *English-Lakota dictionary.* Richmond, Surrey: Curzon Press.

Ingham B (2003). *Lakota: languages of the world/materials 426.* Munich: Lincom Europa.

Parks D R & DeMallie R J (1992). ‘Sioux, Assiniboine and Stoney dialects: a classification.’ *Anthropological Linguistics 34(i–iv)*, 233–256.

Riggs S R (1893). *Dakota grammar texts and ethnography.* Washington: Government Printing office.

Rood D S & Taylor A R (1976). *Lakhota grammar or beginning Lakhota* (vol. 2). Boulder: University of Colorado Lakhota Project.

Lancelot, Claude (1615–1695)

P Swiggers, Katholieke Universiteit Leuven, Leuven, Belgium

Born in Paris, Claude Lancelot entered the Saint Nicolas seminary in 1627, received tonsure in 1628, and became a member of the clergy without being ordained a priest. He studied theology, philosophy, and the classical languages, and also learned two modern Romance languages, viz., Spanish and Italian. Attracted by the doctrine of Saint-Cyran, he became one of the 'solitaires' at Port-Royal (near Paris), where a modern type of (individualized) education of young children was set up in the form of 'small schools' (*petites écoles*).

Lancelot (whose pseudonym was Sir de Trigny) was responsible for the grammatical instruction of the young pupils of these schools, mostly noble youngsters; he wrote manuals for their use, thus publishing a wide range of grammars, all inspired by the then widespread demand of 'methodical order and progression.' In 1644, he published a 'new method' for the teaching of Latin (the fifth edition, of 1656, incorporates important changes brought about by Lancelot's discovery of the works of F. Sanctius, G. J. Vossius, and G. Scioppius), in 1655 a 'new method' for the study of Greek, and in 1660 two 'methodical' manuals for Italian and Spanish. Also in 1660, the anonymous *Grammaire générale et raisonnée* [*de Port-Royal*], appeared; its authors were the theologian and logician A. Arnauld and Lancelot. The work, which offers a didactically oriented discussion on the nature and function of language (seen as a hierarchically-structured system of signs) and the content of grammatical categories (esp. word classes and their *accidentia*), stands at the beginning of the trend of 'general (or philosophical) grammars' in modern times.

Lancelot had a practical, modernist conception of grammar teaching; he preferred to start from the native language of the pupil, and laid stress on a clear typographical presentation, the use of synoptic tables, structural progression in teaching, and intelligent classroom use of high-quality literary texts.

In 1660, the Port-Royal schools were suppressed by royal order, and Lancelot first became a private preceptor, then a Cistercian monk (in the abbey of Saint-Cyran). He was ordained subdeacon in 1673, but left the abbey in order to become a Benedictine in Quimperlé, where he died on April 15, 1695.

See also: Arnauld, Antoine (1612–1694); Sanctius, Franciscus (1523–1600).

Bibliography

Cognet L (1950). *Claude Lancelot, solitaire de Port-Royal.* Paris: Sulliver.
Donetzkoff D (ed.) (2003). *Claude Lancelot: Mémoires touchant la vie de M. de Saint-Cyran.* Paris: Nolin.
Donzé R (1967). *La Grammaire générale et raisonnée de Port-Royal.* Bern: Francke. (2nd edn., 1971).
Gazier C (1932). *Ces Messieurs de Port-Royal.* Paris: Perrin.
Laporte J (1932–1952). *La doctrine de Port-Royal* (4 vols). Paris: P. U. F.
Mertens F J & Swiggers P (1983). 'La *Grammaire générale et raisonnée* de Port-Royal: Notes bibliographiques.' *Historiographia Linguistica 10,* 357–362.
Padley G A (1976–1988). *Grammatical theory in Western Europe, 1500–1700* (3 vols). Cambridge: Cambridge University Press.
Swiggers P (ed.) (1984). *Grammaire et méthode au XVII^{e} siècle.* Leuven: Peeters.
Swiggers P (1987). 'La sémiotique de Port-Royal: du savoir au vouloir(-dire).' *Semiotica 66,* 331–344.
Swiggers P (1989). 'La *Grammaire* de Port-Royal et le parallélisme logico-grammatical.' *Orbis 33,* 29–56.
Weaver E (1978). *The evolution of the reform of Port-Royal.* Paris: Beauchesne.

Landau-Kleffner Syndrome

S Rémillard and H Cohen, Université du Québec à Montréal, Montréal, Canada
M Plaza, CNRS and Université Paris V René Desceartes, France

Landau-Kleffner syndrome (LKS), also called acquired epileptic aphasia, was first described by Landau and Kleffner in 1957. This syndrome is a rare childhood neurological disorder characterized by the sudden or gradual loss of receptive and expressive language skills and paroxysmal electroencephalographic (EEG) abnormalities. The EEG abnormalities are characterized by spikes, sharp waves, or spike–wave discharges usually recorded bilaterally and mainly over the temporoparietal (perisylvian) regions (Tuchman and Rapin, 2002). However, in 15% of cases, the epileptiform discharges are reported to be unilateral and

Lancelot, Claude (1615–1695)

P Swiggers, Katholieke Universiteit Leuven, Leuven, Belgium

Born in Paris, Claude Lancelot entered the Saint Nicolas seminary in 1627, received tonsure in 1628, and became a member of the clergy without being ordained a priest. He studied theology, philosophy, and the classical languages, and also learned two modern Romance languages, viz., Spanish and Italian. Attracted by the doctrine of Saint-Cyran, he became one of the 'solitaires' at Port-Royal (near Paris), where a modern type of (individualized) education of young children was set up in the form of 'small schools' (*petites écoles*).

Lancelot (whose pseudonym was Sir de Trigny) was responsible for the grammatical instruction of the young pupils of these schools, mostly noble youngsters; he wrote manuals for their use, thus publishing a wide range of grammars, all inspired by the then widespread demand of 'methodical order and progression.' In 1644, he published a 'new method' for the teaching of Latin (the fifth edition, of 1656, incorporates important changes brought about by Lancelot's discovery of the works of F. Sanctius, G. J. Vossius, and G. Scioppius), in 1655 a 'new method' for the study of Greek, and in 1660 two 'methodical' manuals for Italian and Spanish. Also in 1660, the anonymous *Grammaire générale et raisonnée* [*de Port-Royal*], appeared; its authors were the theologian and logician A. Arnauld and Lancelot. The work, which offers a didactically oriented discussion on the nature and function of language (seen as a hierarchically-structured system of signs) and the content of grammatical categories (esp. word classes and their *accidentia*), stands at the beginning of the trend of 'general (or philosophical) grammars' in modern times.

Lancelot had a practical, modernist conception of grammar teaching; he preferred to start from the native language of the pupil, and laid stress on a clear typographical presentation, the use of synoptic tables, structural progression in teaching, and intelligent classroom use of high-quality literary texts.

In 1660, the Port-Royal schools were suppressed by royal order, and Lancelot first became a private preceptor, then a Cistercian monk (in the abbey of Saint-Cyran). He was ordained subdeacon in 1673, but left the abbey in order to become a Benedictine in Quimperlé, where he died on April 15, 1695.

See also: Arnauld, Antoine (1612–1694); Sanctius, Franciscus (1523–1600).

Bibliography

Cognet L (1950). *Claude Lancelot, solitaire de Port-Royal.* Paris: Sulliver.

Donetzkoff D (ed.) (2003). *Claude Lancelot: Mémoires touchant la vie de M. de Saint-Cyran.* Paris: Nolin.

Donzé R (1967). *La Grammaire générale et raisonnée de Port-Royal.* Bern: Francke. (2nd edn., 1971).

Gazier C (1932). *Ces Messieurs de Port-Royal.* Paris: Perrin.

Laporte J (1932–1952). *La doctrine de Port-Royal* (4 vols). Paris: P. U. F.

Mertens F J & Swiggers P (1983). 'La *Grammaire générale et raisonnée* de Port-Royal: Notes bibliographiques.' *Historiographia Linguistica 10,* 357–362.

Padley G A (1976–1988). *Grammatical theory in Western Europe, 1500–1700* (3 vols). Cambridge: Cambridge University Press.

Swiggers P (ed.) (1984). *Grammaire et méthode au XVII*[e] *siècle.* Leuven: Peeters.

Swiggers P (1987). 'La sémiotique de Port-Royal: du savoir au vouloir(-dire).' *Semiotica 66,* 331–344.

Swiggers P (1989). 'La *Grammaire* de Port-Royal et le parallélisme logico-grammatical.' *Orbis 33,* 29–56.

Weaver E (1978). *The evolution of the reform of Port-Royal.* Paris: Beauchesne.

Landau-Kleffner Syndrome

S Rémillard and H Cohen, Université du Québec à Montréal, Montréal, Canada
M Plaza, CNRS and Université Paris V René Desceartes, France

Landau-Kleffner syndrome (LKS), also called acquired epileptic aphasia, was first described by Landau and Kleffner in 1957. This syndrome is a rare childhood neurological disorder characterized by the sudden or gradual loss of receptive and expressive language skills and paroxysmal electroencephalographic (EEG) abnormalities. The EEG abnormalities are characterized by spikes, sharp waves, or spike–wave discharges usually recorded bilaterally and mainly over the temporoparietal (perisylvian) regions (Tuchman and Rapin, 2002). However, in 15% of cases, the epileptiform discharges are reported to be unilateral and

show temporal lobe predominance. The major characteristic of LKS is the presence of continuous spike-and-waves EEG discharges during 85% of the slow sleep period (Hirsh *et al.*, 1990). All six children with LKS in the Lewine *et al.* (1997) study showed continuous spike-and-waves during slow sleep (CSWS). Onset of aphasia in LKS typically occurs between 3 and 9 years of age. There is sex predominance with a male:-female ratio of approximately 2:1. Although language regression is the key feature of clinically typical LKS, cognitive and behavioral dysfunctions such as motor hyperactivity, decreased attention, aggressiveness, and social withdrawal can also accompany this disorder (Pearl *et al.*, 2001).

Clinical epileptic seizures are seen in many of the affected children either before or after the onset of aphasia (Beaumanoir, 1992). Although 70–80% of patients show seizures (Hirsch *et al.*, 1990), epilepsy is not an essential criterion for the diagnosis of LKS, in contrast to the EEG abnormalities that are present in all LKS cases (Beaumanoir, 1992; Sobel *et al.*, 2000). The usual types of epilepsy observed in LKS may be simple partial motor, generalized tonic–clonic, or atypical absence seizures (Beaumanoir, 1992; Metz-Lutz *et al.*, 1999; Morrell *et al.*, 1995). Myoclonic and tonic seizures are uncommon (Deonna, 1991). In general, the seizures are easily controlled with anticonvulsive medication and, like the EEG abnormalities, remit by the end of adolescence (Beaumanoir, 1992; Hirsch *et al.*, 1990). The etiology of LKS remains unknown and may be due to diverse causes. Some cases of LKS have been associated with toxoplasmosis (Michalowicz *et al.*, 1989), inflammatory demyelinating illness (Perniola *et al.*, 1993), temporal lobe tumor (Nass *et al.*, 1993; Solomon *et al.*, 1993), and neurocysticercosis (Otero *et al.*, 1989).

There is no evidence of anatomical brain lesions. Furthermore, the results of computed tomographic scan (Guerreiro *et al.*, 1996; Hirsch *et al.*, 1990) and magnetic resonance imaging (Hirsch *et al.*, 1990) appear normal. However, single-photon-emission computed tomography and positron-emission tomography (PET) studies have shown unilateral left temporoparietal or bitemporal abnormalities as manifested by increased glucose utilization (Guerreiro *et al.*, 1996; Rintahaka *et al.*, 1995).

Linguistic Aspects

The first manifestation of the language disturbance in LKS is an apparent 'word deafness' or verbal auditory agnosia, occurring in a child with previously normal language development. In many cases, this auditory agnosia extends to familiar sounds including bells, whistles, or a ringing phone (Pearl *et al.*, 2001). Generally, it is the first symptom preceding the epileptic manifestations and expressive language dysfunction. At first, these children may appear to be deaf. However, hearing tests, such as brainstem auditory evoked responses and conventional audiologic assessment, usually reveal normal conduction of acoustic stimuli from peripheral level to the primary auditory cerebral cortex (Pearl *et al.*, 2001). Thus, some authors consider this auditory recognition impairment as a specific deficit in higher phonological decoding processes (e.g., Korkman *et al.*, 1998), and others focus on the dissociation between auditory and phonological processing (Plaza *et al.*, 2001).

The second manifestation of language dysfunction is characterized by an impaired expressive communication. There is a progressive decline in the vocabulary of the affected child accompanied by difficulties in articulation and telegraphic speech disturbances (Metz-Lutz *et al.*, 1999). The expressive language dysfunction gradually increases and leads to frank mutism (Beaumanoir, 1992; Metz-Lutz *et al.*, 1999; Morrell *et al.*, 1995; Pearl *et al.*, 2001). Paraphasias and neologisms are often observed before the complete loss of expressive language (Metz-Lutz *et al.*, 1999). Agrammatism, echolalia, anomia, and phonetic distortion have also been reported in some cases of LKS (Paquier *et al.*, 1992).

Language disturbance may fluctuate significantly during the course of LKS, with remissions and exacerbations (Lanzi *et al.*, 1994; Pearl *et al.*, 2001), but total recovery is uncommon (Santos *et al.*, 2002). Although spontaneous recovery has been reported (Santos *et al.*, 2002; Deonna *et al.*, 1989), the majority of children with LKS are left with significant language deficits (e.g., severe dyslexia, severe to moderate expressive aphasia, or global aphasia; Bishop, 1985; Deonna *et al.*, 1989). Morrell *et al.* (1995) have proposed that excessive epileptic discharge during the childhood period of language development, in children with LKS, results in functionally inappropriate synaptogenesis and then deficient synaptic pruning in the same regions subserving language function. Thus, nonfunctional synaptic contacts remain in the language cortex even after the epileptic process has disappeared, leaving residual language deficits. Older children, who have learned to read and write, may lose these abilities in the course of the disease. However, a partial retention of writing skills may suggest a better prognosis in the re-education phase (Beaumanoir, 1992).

Originally, Landau and Kleffner (1957) reported a correlation between the degree of language dysfunction and the severity of the EEG abnormalities. However, the relationship between EEG and aphasia in LKS is not clear and still controversial. Some authors

have shown considerable speech recovery in patients with the disappearance of EEG abnormalities when successfully treated with anticonvulsant medication (Lanzi *et al.*, 1994; Morrell *et al.*, 1995; Sobel *et al.*, 2000). Others have shown no association between linguistic abilities and EEG abnormalities. Improvement or worsening of aphasia occurs independently of the EEG abnormalities (Deonna, 1991). In contrast to the hypothesis initially proposed by Landau and Kleffner, Holmes and Riviello (2001) suggested that the epileptic discharges are the epiphenomena of the underlying pathology of cerebral speech areas, rather than the cause of the aphasia. Metz-Lutz *et al.* (1999) associated the persisting verbal impairments in LKS patients with a phonological short-term memory deficit. A PET study has shown less activation in the right and left superior temporal gyri of adults who recovered from LKS during a task of serial words recall (Majerus *et al.*, 2003).

Treatment and Prognosis

As most authors have emphasized, the seizure disorder and the EEG abnormalities tend to remit by the age of 12 or 13, whereas language dysfunction persists into adulthood to variable degrees (Deonna, 1991; Metz-Lutz *et al.*, 1999; Morrell *et al.*, 1995). However, case studies report variable prognosis concerning the evolution of language dysfunction in LKS children. Few children recover completely from their aphasia (Santos *et al.*, 2002), whereas others show complete absence of speech (Deonna, 1991). Nevertheless, the majority of LKS patients show partial recovery of linguistic abilities. A late outcome study investigating late auditory evoked potentials over the temporal areas previously involved by epileptic discharges has revealed a permanent dysfunction in the associative auditory cortex, characterized by a unilateral extinction on dichotic listening tests (Wioland *et al.*, 2001). To compensate for language deprivation during the active period of LKS, rehabilitation strategies, such as sign language or cued speech, are particularly encouraged (Gordon, 2004; Metz-Lutz *et al.*, 1999). These compensatory strategies provide an alternative means of communication that may facilitate recovery of speech and language.

According to Bishop (1985), an important factor negatively influencing language prognosis is the early age of onset (i.e., before the age of 5 years). Another critical factor influencing prognosis is the duration of the EEG abnormalities and the period of CSWS. As suggested by Morell *et al.* (1995), the long-term effect of the epileptiform discharges on brain cells of a cortical area results in the disruption of normal axonal pruning. Grote *et al.*'s (1999) study, in which the patients with the shortest duration of illness were most improved, provides good support for this hypothesis.

One question is how reading and spelling could be acquired by children with LKS. A recent case study of a girl (subject M.M.) followed for five years and after recovery (Plaza, 2004), focused on written language, which M.M. acquired in spite of persistent but heterogeneous phonological impairment (efficient retrieval of phonological information from long-term memory, restriction of working memory and slow processing in phonological awareness tasks). M.M. developed above-average reading and spelling skills, notably for real words, and gradually mastered pseudoword reading. However, she remained impaired in pseudoword spelling. The fact that M.M.'s performance on pseudoword tasks was poor when the stimuli were presented auditorily, and more efficient when they were presented visually, suggested that the observed phonological impairment was (1) due to a deficit at the level of cortical auditory processing, and (2) related to the temporal lobe dysfunction observed through the long-lasting right channel extinction in dichotic listening tasks.

The treatment of LKS aims at eliminating or reducing the epileptiform activity early in the course of the disease and includes anticonvulsants (Marescaux *et al.*, 1990), immunoglobulin (Mikati and Saab, 2000), and corticosteroid (Tsuru *et al.*, 2000) therapies. However, the outcome on language following pharmacological treatment remains variable and unpredictable. In the case of intractable epilepsy, multiple subpial transection surgery, a more radical and controversial therapeutic method, has met with successful results (Grote *et al.*, 1999; Morell *et al.*, 1995).

In summary, LKS provides an important model of the way in which epilepsy can affect language and cognition. This language disorder reinforces the hypothesis that focal epileptic activity occurring during a critical period of brain maturation may result in a permanent cerebral dysfunction which nevertheless can be compensated for by cognitive strategies.

See also: Aphasia Syndromes; Autism and Asperger Syndrome: A Spectrum of Disability; Developmental Relationship between Language and Cognition.

Bibliography

Beaumanoir A (1992). 'The Landau-Kleffner syndrome.' In Roger J, Beureau M, Dravet C, Dreifuss F E, Perret A & Wolf P (eds.) *Epileptic syndromes in infancy, childhood and adolescence,* 2nd edn. London: John Libbey. 231–243.

Bishop D V M (1985). 'Age of onset and outcome in "acquired aphasia with convulsive disorder."' *Developmental Medicine and Child Neurology 27,* 705–712.

Deonna T W (1991). 'Acquired epileptiform aphasia in children (Landau-Kleffner syndrome).' *Journal of Clinical Neurophysiology 8,* 288–298.

Deonna T, Peter C & Ziegel A L (1989). 'Adult follow-up of the acquired aphasia–epilepsy syndrome in childhood. Report of 7 cases.' *Neuropediatrics 20,* 132–138.

Gordon N (2004). 'The neurology of sign language.' *Brain & Development 26,* 146–150.

Grote C L, Van-Slyke P & Hoeppner J-A B (1999). 'Language outcome following multiple subpial transection for Landau-Kleffner syndrome.' *Brain 122, 561–566.*

Guerreiro M M, Camargo E E, Kato M, Menezes Netto Jr, Silva E A, Scotoni A E, Silveira D C & Guerreiro C A (1996). 'Brain single photon emission computed tomography imaging in Landau-Kleffner syndrome.' *Epilepsia 37,* 60–67.

Hirsch E, Marescaux C, Maquet P, Metz-Lutz M N, Kiesmann M, Salomon E, Franck G & Kurtz D (1990). 'Landau-Kleffner syndrome: A clinical and EEG study of five cases.' *Epilepsia 31,* 756–767.

Holmes G L & Riviello J J Jr (2001). 'Treatment of childhood idiopathic language deterioration with valproate.' *Epilepsy & Behavior 2,* 272–276.

Korkman M, Granström M-L, Appleqvist K & Liukkomen E (1998). 'Neuropsychological characteristics of five children with the Landau-Kleffner syndrome: Dissociation of auditory and phonolical discrimination.' *Journal of International Neuropsychological Society 4, 566–575.*

Landau W M & Kleffner F R (1957). 'Syndrome of acquired aphasia with convulsive disorder in children.' *Neurology 7,* 523–530.

Lanzi G, Veggiotti P, Conte S, Partesana E & Resi C (1994). 'A correlated fluctuation of language and EEG abnormalities in a case of the Landau-Kleffner syndrome.' *Brain & Development 16,* 329–334.

Lewine J, Chez M, Andrews R, Devinski O, Morrell F & Smith M (1997). 'Neuromagnetic evaluation of children with acquired developmental aphasia.' *Epilepsia 38,* 128.

Majerus S, Laureys S, Collette F, Del Fiore G, Degueldre C, Luxen A, Van der Linden M, Maquet P & Metz-Lutz M-N (2003). 'Phonological short-term memory networks following recovery from Landau and Kleffner syndrome.' *Human Brain Mapping 19,* 133–144.

Marescaux C, Hirsch E, Finck P, Maquet P, Schlumberger E, Sellal F, Metz-Lutz M N, Alembik Y, Salmon E & Franck G (1990). 'Landau-Kleffner syndrome: A pharmacologic study of five cases.' *Epilepsia 31,* 768–777.

Metz-Lutz M-N, Seegmuller C, Kleitz C, de Saint Martin A, Hirsch E & Merescaux C (1999). 'Landau-Kleffner syndrome: A rare childhood epileptic aphasia.' *Journal of Neurolinguistics 12,* 167–179.

Michalowiscz R, Joswiak S, Ignatowicz R & Szwabowska-Oreszko E (1989). 'Landau-Kleffner syndrome: Epileptic aphasia in children: Possible role of *Toxoplasma gondi* infection.' *Acta Paediatrica Hungarica 29,* 337–342.

Mikati M & Saab R (2000). 'Successful use of intravenous immunoglobulin as initial monotherapy in Landau-Kleffner syndrome.' *Epilepsia 41,* 880–886.

Morrell F, Whistler W W, Smith M C, Hoeppner T J, de Toledo-Morrell L, Pierre-Louis S J, Kanner A M, Buelow J M, Ristanovic R & Bergen D (1995). 'Landau-Kleffner syndrome: Treatment with subpial intracortical transection.' *Brain 118,* 1529–1546.

Nass R, Heier L & Walker R (1993). 'Landau-Kleffner syndrome: Temporal lobe tumor resection results in good outcome.' *Pediatric Neurology 9,* 303–305.

Otero E, Cordova S, Diaz F, Garcia-Teruel I & DelBrutto O H (1989). 'Acquired epileptic aphasia (the Landau-Kleffner syndrome) due to neurocysticercosis.' *Epilepsia 30, 569–572.*

Paquier P F, Van Dongen H R & Loonen C B (1992). 'The Landau-Kleffner syndrome or "acquired aphasia with convulsive disorder." Long-term follow-up of six children and a review of recent literature.' *Archives of Neurology 49,* 354–359.

Pearl P L, Carrazana E J & Holmes G L (2001). 'The Landau-Kleffner syndrome.' *Epilepsy Currents 1,* 39–45.

Perniola T, Magari L, Buttiglione M, Andreula C, Simone I L & Santostasi R (1993). 'A case of Landau-Kleffner syndrome secondary to inflammatory demyelinating disease.' *Epilepsia 39, 551–556.*

Plaza M (2004). Reading and spelling skills in a child with a Landau–Kleffner Syndrome. *Second Lisbon Meeting on Language Acquisition.* Lisbon, Portugal, 1–4 June.

Plaza M, Rigoard M-T, Chevrie-Muller C, Cohen H & Picard A (2001). 'Short term memory impairment and unilateral dichotic listening extinction in a child with Landau–Kleffner Syndrome: Auditory or phonological disorder?' *Brain and Cognition 46,* 235–240.

Rintahaka P J, Chugani H T & Sankar R (1995). 'Landau-Kleffner syndrome with continuous spikes and waves during slow-wave sleep.' *Journal of Child Neurology 10,* 127–133.

Santos L H, Antoniuk S A, Rodrigues M, Bruno S & Bruck I (2002). 'Landau-Kleffner syndrome.' *Arquivos de Neuropsiquiatria 60,* 239–241.

Sobel D F, Aung M, Otsubo H & Smith M C (2000). 'Magnetoencephalography in children with Landau-Kleffner syndrome and acquired epileptic aphasia.' *American Journal of Neuroradiology 21,* 301–307.

Solomon G E, Carson D, Paulakis S, Fraser R & Labar D (1993). 'Intracranial EEG monitoring in Landau-Kleffner syndrome associated with temporal lobe astrocytoma.' *Epilepsia 34,* 557–560.

Tsuru T, Mori M, Mizuguchi M & Momoi M (2000). 'Effects of high-dose intravenous corticosteroid therapy in Landau-Kleffner syndrome.' *Pediatric Neurology 22,* 145–147.

Tuchman R & Rapin I (2002). 'Epilepsy in autism.' *Lancet Neurology 1,* 352–358.

Wioland N, Rudolf G & Metz-Lutz M-N (2001). 'Electrophysiological evidence of persisting unilateral auditory cortex dysfunction in the late outcome of Landau and Kleffner syndrome.' *Clinical Neurophysiology 112,* 319–323.

Langacker, Ronald W. (b. 1942)

J-M Fortis, CNRS, University of Paris, Paris, France

Ronald W. Langacker was born on December 27, 1942 in Fond du Lac (Wisconsin, USA). He graduated from the University of Illinois in 1966 before moving to the University of California at San Diego, where he now occupies a chair of linguistics. At the beginning of his career he did fieldwork, notably on Luiseño and Papago (O'odham) (two Uto-Aztecan languages). The outcome was a four-volume survey on Uto-Aztecan languages for which Langacker himself wrote a general introduction and assumed editorial control. His contribution to Uto-Aztecan linguistics also includes a monograph bearing on reciprocal/reflexive predicates and their relation to passives and impersonals.

In the mid-1970s, Langacker laid the foundations of a theoretical framework called 'space grammar.' This designation was apparently motivated by the idea that linguistic facts, like vision, involve a perspective on situations from a vantage point centered on ego. This theory was notably applied to the analysis of English epistemic modals, whose order was correlated with a layered structure of conceptualizations.

The theory was renamed 'cognitive grammar' in the 1980s (for an introduction, see Langacker, 1982 and 1986). It is cognitive in several respects: (1) grammatical notions are given definitions that appeal to properties of perceptual representations, e.g., the grammatical subject and object are defined in terms of figure/ground organization of conceptualized situations; (2) linguistic structures are formed and categorized by cognitive processes that are shared by other cognitive systems. Thus, the ability to form phrase schemas rests on the capacity for abstracting commonalities from inputs instantiating any more abstract structure; since this is a psychological process, schemas of various degrees of generality may proliferate (cognitive grammar seeks no economical representation of linguistic structures); (3) linguistic representations are structured like other representations; for instance, membership in a category is not an all-or-nothing affair but a matter of degrees (typicality effects); (4) some linguistic phenomena reflect the operation of 'imagery,' i.e., the ability to form alternate construals of the same situation. Thus, ergative and accusative systems are claimed to reflect 'conceptual paths' originating from, respectively, the undergoer of a process and the initiator of an action; (5) the conceptualization of reality is based on cognitive models of the world, in particular, a naive physical theory of causal interactions.

Cognitive grammar has become a major theoretical trend within cognitive linguistics. Further, after the creation of the journal *Cognitive Linguistics* and of the International Cognitive Linguistics Association (in 1989), cognitive grammar has gained an international influence.

See also: Cognitive Grammar; Uto-Aztecan Languages.

Bibliography

Langacker R W (1977). *Studies in Uto-Aztecan grammar, vol. 1: an overview of Uto-Aztecan grammar.* Dallas: Summer Institute of Linguistics and University of Texas at Arlington.

Langacker R W (1982). 'The English passive.' *Language 58,* 22–80 [reprinted in R W Langacker (1991), 101–147].

Langacker R W (1986). An introduction to cognitive grammar.' *Cognitive Science 10,* 1–40 [reprinted in Langacker R W (1991), 1–32].

Langacker R W (1987–1991). *Foundations of cognitive grammar* (vols 1–2). Stanford: Stanford University Press.

Langacker R W (1991). *Concept, image and symbol.* Berlin & New York: Mouton de Gruyter. 1–32.

Language Acquisition Research Methods

E Kidd, University of Manchester, Manchester, UK

As a subdiscipline of psycholinguistics the broad aim of language acquisition research is to understand the architecture and development of the language faculty. In fulfilling this aim, child language researchers are faced with a problem pervasive in all disciplines of science: data collection relies on observation, which is indirect and theory dependent. To overcome this problem researchers use a vast array of research methodologies that provide convergent evidence for or against a current position or theory. This article surveys the types of methods that are used in the field.

Langacker, Ronald W. (b. 1942)

J-M Fortis, CNRS, University of Paris, Paris, France

Ronald W. Langacker was born on December 27, 1942 in Fond du Lac (Wisconsin, USA). He graduated from the University of Illinois in 1966 before moving to the University of California at San Diego, where he now occupies a chair of linguistics. At the beginning of his career he did fieldwork, notably on Luiseño and Papago (O'odham) (two Uto-Aztecan languages). The outcome was a four-volume survey on Uto-Aztecan languages for which Langacker himself wrote a general introduction and assumed editorial control. His contribution to Uto-Aztecan linguistics also includes a monograph bearing on reciprocal/reflexive predicates and their relation to passives and impersonals.

In the mid-1970s, Langacker laid the foundations of a theoretical framework called 'space grammar.' This designation was apparently motivated by the idea that linguistic facts, like vision, involve a perspective on situations from a vantage point centered on ego. This theory was notably applied to the analysis of English epistemic modals, whose order was correlated with a layered structure of conceptualizations.

The theory was renamed 'cognitive grammar' in the 1980s (for an introduction, see Langacker, 1982 and 1986). It is cognitive in several respects: (1) grammatical notions are given definitions that appeal to properties of perceptual representations, e.g., the grammatical subject and object are defined in terms of figure/ground organization of conceptualized situations; (2) linguistic structures are formed and categorized by cognitive processes that are shared by other cognitive systems. Thus, the ability to form phrase schemas rests on the capacity for abstracting commonalities from inputs instantiating any more abstract structure; since this is a psychological process, schemas of various degrees of generality may proliferate (cognitive grammar seeks no economical representation of linguistic structures); (3) linguistic representations are structured like other representations; for instance, membership in a category is not an all-or-nothing affair but a matter of degrees (typicality effects); (4) some linguistic phenomena reflect the operation of 'imagery,' i.e., the ability to form alternate construals of the same situation. Thus, ergative and accusative systems are claimed to reflect 'conceptual paths' originating from, respectively, the undergoer of a process and the initiator of an action; (5) the conceptualization of reality is based on cognitive models of the world, in particular, a naive physical theory of causal interactions.

Cognitive grammar has become a major theoretical trend within cognitive linguistics. Further, after the creation of the journal *Cognitive Linguistics* and of the International Cognitive Linguistics Association (in 1989), cognitive grammar has gained an international influence.

See also: Cognitive Grammar; Uto-Aztecan Languages.

Bibliography

Langacker R W (1977). *Studies in Uto-Aztecan grammar, vol. 1: an overview of Uto-Aztecan grammar.* Dallas: Summer Institute of Linguistics and University of Texas at Arlington.

Langacker R W (1982). 'The English passive.' *Language 58,* 22–80 [reprinted in R W Langacker (1991), 101–147].

Langacker R W (1986). An introduction to cognitive grammar.' *Cognitive Science 10,* 1–40 [reprinted in Langacker R W (1991), 1–32].

Langacker R W (1987–1991). *Foundations of cognitive grammar* (vols 1–2). Stanford: Stanford University Press.

Langacker R W (1991). *Concept, image and symbol.* Berlin & New York: Mouton de Gruyter. 1–32.

Language Acquisition Research Methods

E Kidd, University of Manchester, Manchester, UK

As a subdiscipline of psycholinguistics the broad aim of language acquisition research is to understand the architecture and development of the language faculty. In fulfilling this aim, child language researchers are faced with a problem pervasive in all disciplines of science: data collection relies on observation, which is indirect and theory dependent. To overcome this problem researchers use a vast array of research methodologies that provide convergent evidence for or against a current position or theory. This article surveys the types of methods that are used in the field.

Naturalistic Speech Data

Researchers have collected naturalistic speech data from children as far back as Darwin (1877). This was done initially using diary studies, whereby researchers recorded the child's speech in a logbook. The focus was on phonological, lexical, or syntactic development. The collection of naturalistic speech data was revolutionized by the technological advances that enabled children's speech to be recorded. This was particularly the case for researchers studying phonological development, where recording enabled more accurate categorization of speech patterns. Brown's (1973) longitudinal study of three English-speaking children set the standard for the discipline. The construction of an electronic archive of children's transcribed naturalistic speech, the Child Language Data Exchange System (CHILDES; see MacWhinney, 2000), has enabled researchers from all over the world to share and use data to answer a multitude of research questions.

Naturalistic speech data has numerous advantages. In particular, longitudinal data can be useful for identifying general developmental trends. This is particularly useful for languages that have not been studied extensively (Demuth, 1996). The disadvantage of naturalistic speech data is that such data can underestimate children's knowledge, since children do not always say what they know. Furthermore, low-frequency linguistic patterns are unlikely to be recorded, since practical limitations in data collection mean that naturalistic studies record only a small percentage of children's total speech (Tomasello and Stahl, 2004). It is for these reasons that experimental data are useful for determining children's linguistic knowledge.

Experimental Approaches

Experimental approaches to the study of child language probe either children's production or comprehension. The choice to study production or comprehension depends on one's research question, as does the particular method used. Since children are generally assumed to understand more than they say (Shipley *et al.*, 1969), comprehension tasks enable researchers to test the limits of children's linguistic knowledge. On the other hand, children's productive use of language in experimental settings can potentially enable researchers to make strong claims about the abstractness of knowledge, for instance, knowledge of syntactic and phrasal categories, or discourse competence, such as when children correctly use various linguistic devices to introduce and resume reference to entities in discourse.

Production Tasks

There are numerous ways to elicit language from children in experimental settings. Since children are not aware of the linguistic target of interest, such studies must be carefully designed to restrict the range of linguistic options that children can produce. For instance, Berko's (1958) 'Wug' test investigated children's morphosyntactic knowledge by creating an appropriate sentential context. Variants of this task, which make use of novel words, are still being used in present-day research. Other researchers have been successful in eliciting various syntactic forms from children (Crain and Thornton, 1998). Such studies show that target forms can be elicited when the correct aspects of the environment are manipulated and when children are probed with questions that constrain potential answers. Syntactic priming, where children are primed with a target structure prior to being required to produce language, is another method that increases the chance that a particular structure will be produced. Another production method that is useful for investigating children's syntactic knowledge is elicited imitation. In this method, children repeat sentences spoken to them by an experimenter. Exact repetitions suggest that children are competent with the structure under investigation. Children's spontaneous changes to the test structure can provide some insights into the child's analysis and interpretation of the adult model.

A less constrained method of eliciting language from children is to collect narrative data. One particularly useful method has been to ask children to describe picture books in which a protagonist engages in an activity attractive to children. One book on which there is a substantial amount of data is Mayer's (1969) *Frog, Where Are You?* (see Berman and Slobin, 1994). Narrative data provide useful insights into children's knowledge of discourse devices, such as referent introduction and anaphora, and into the discourse bases of grammatical constructions that encode both new and old information, such as many types of complex sentences. Referential communication tasks (see Lloyd *et al.*, 1998) enable one to investigate children's communicative competence and, in particular, their understanding of conversational implicatures. This method therefore indicates children's understanding of a conversational partner's discourse knowledge.

Comprehension Tasks

Comprehension tasks provide the experimenter with a greater degree of control over the outcome of an experiment than do production tasks. This is largely because children's linguistic knowledge is inferred

from certain behaviors; for instance, their preference for one interpretation or referent over another, or their 'yes/no' judgment of the grammatical acceptability of a sentence. Since children's knowledge is inferred from their behavior, researchers must be careful to design studies with exclusionary criteria that identify children who are using non-linguistic strategies in the task, such as always providing affirmative judgments of sentences on every trial of an experiment investigating children's grammaticality judgments.

Tasks that require children to select a referent or interpretation from a set are useful because they can be used to test children's knowledge at either the lexical or the sentential level. Either pictures or real objects can be used in word learning experiments; experiments conducted at the sentential level, where a number of interpretations are provided, generally use pictures. The preferential looking and preferential listening paradigms provide alternative dependent measures; they use the continuous variables of looking or listening time to infer children's knowledge (Fernald, 1985; Golinkoff *et al.*, 1987; Jusczyk, 1997). These tasks are useful because they can be used with very young infants, since tasks demands are greatly reduced. Preferential listening (or 'head preference') is most often used to investigate children's speech perception. This technique can, for instance, probe children's sensitivity to the phonotactic regularities and phonemic contrasts of their input language. It can also be used to investigate young children's sensitivity to prosody. One disadvantage of preferential looking and listening is that they yield rather noisy individual data and therefore require one to test enough participants to ensure that stable behavioral patterns are observed.

One popular comprehension method is the act-out task, where children are provided with toy figurines that act as referents for participants in test sentences. The act-out methodology is flexible enough to be used for variety of syntactic phenomena. One disadvantage is that the task imposes considerable demands on children, who must process the test sentence and translate their interpretations onto the toy referents provided (Hamburger and Crain, 1987). This problem can be partially alleviated by providing discourse contexts to test sentences. Another set of methods that probe children's comprehension are judgment tasks. These require children to provide yes/no judgments about the acceptability of linguistic stimuli. When used with very young children, these tasks usually involve a puppet who, the child is informed, "hasn't learned to talk properly" (Lloyd and Donaldson, 1976). The child is then required to judge whether the puppet has used the correct language in a given situation. Metalinguistic tasks, such as grammaticality judgments, which are generally conducted with older children (Bialystok, 1986; Wufleck, 1993), simply require a yes/no response. These tasks enable the researcher to determine which aspects of language children find acceptable, and are particularly useful for probing syntactic knowledge.

One final set of comprehension tasks are those that yield 'on-line' reaction time data. There are many on-line techniques, but the best are those that do not require children to read. Some techniques traditionally used to test adults have been developed for use with children, such as mispronunciation detection, cross-modal priming, word monitoring (see McKee, 1996), and self-paced listening (Booth *et al.*, 2000). These 'button-pressing' techniques require children to make decisions about different aspects of language (e.g., lexical decisions, judgments, and comprehension questions). Finally, the eye-tracking technique records children's eye movements to toys or pictures as they hear a test sentence (Trueswell *et al.*, 1999). One clear advantage of on-line tasks is that they explicitly test children's parsing systems, which can in turn suggest the aspects of language that children are most sensitive to as they hear it.

Comparative Methods

Comparative studies in child language research take one of two forms: comparative studies of methods and cross-linguistic comparisons. The former are generally conducted within a given language and provide convergent evidence for the psychological reality of an observed experimental phenomenon (e.g., a comparison of naturalistic and experimental data). The latter have been used extensively in acquisition research to identify universal features of the acquisition process. This approach has been taken by Slobin *et al.* (1985, 1992, 1997). Extensive cross-linguistic data have also been collected on children's comprehension strategies (MacWhinney and Bates, 1989) and children's narrative (Berman and Slobin, 1994). Furthermore, ethnographic research investigates how socialization influences acquisition, recognizing that the child learns language in a rich social context (Ochs and Schieffelin, 1995).

Computational Approaches

Data collected from human participants enable researchers to hypothesize acquisition mechanisms and suggest potential architectural aspects of the language faculty. However, as noted previously, these phenomena are not directly observable. Computational approaches to language acquisition enable

researchers to build hypothesized architectures that are augmented by learning mechanisms and compare the output of these models to data from human participants. Since the manner in which computational models learn language is more transparent than that of the child, computational research provides valuable insights into the acquisition process. There are many different families of computational architectures. One popular computational approach in child language research has been connectionism (Rumelhardt and McClelland, 1986; Elman *et al*., 1996), but other symbolic approaches have made equally significant contributions (Brent, 1997).

Standardized Language Tests

Standardized language tests provide population norming data from which individual children's linguistic knowledge can be estimated relative to their age. This information is useful for a number of reasons. First, many comprehensive tests can be used to diagnose language impairment. Second, standardized language tests can be used to estimate the language level of special populations so that individuals can be matched with normally developing children of the same language ability for purposes of comparison.

Which Method Is Best?

Every method has its advantages and disadvantages. In general, a researcher's choice of method depends on the research question and the age range of the children who will participate in the study. For instance, most on-line methods are unsuitable for children under the age of 3 years, who generally do not understand the requirements of the task. Conversely, preference techniques such as preferential looking and listening are generally suitable only for infants, as older children will quickly become bored with the task. Certain methods are more suitable than others for certain linguistic phenomena. For instance, if studied experimentally, children's syntactic knowledge is best investigated using techniques such as act-out, elicited imitation, or elicited production, where it is possible to manipulate the language children hear or the contexts that lead to the use of different grammatical structures. The acquisition of concrete lexical categories such as nouns is best investigated using preference methods such as picture or referent selection and preferential looking.

The choice of the method will affect the choice of statistics one uses to analyze the data. Cross-sectional designs, which test a number of different age groups, will generally require parametric or nonparametric statistics, depending on the dependent measure. It is possible to analyze longitudinal naturalistic speech data using advanced statistical techniques such as regression or time-series modeling, provided one has enough data points and the assumptions of the statistical technique are not violated. However, sometimes a research question can be answered using simple statistical conventions, such as reporting mean values for variables of interest.

Conclusion

There are numerous methods available for investigating child language acquisition. Choosing a particular method will depend on a number of factors, including the research question, the age of the children to be tested, and other practical considerations, such as expense. The range of methods available enables researchers to approach problems from a number of angles, which can serve to strengthen theoretical claims if convergent evidence is yielded using different methodologies.

See also: CHILDES Database; Cross-Linguistic Comparative Approaches to Language Acquisition.

Bibliography

Berko J (1958). 'The child's learning of English morphology.' *Word 14,* 150–177.
Berman R A & Slobin D I (eds.) (1994). *Relating events in narrative: A cross-linguistic developmental study.* Hillsdale, NJ: Erlbaum.
Bialystok E (1986). 'Factors in the growth of linguistic awareness.' *Child Development 57,* 498–510.
Booth J R, MacWhinney B & Harasaki Y (2000). 'Developmental differences in visual and auditory processing of complex sentences.' *Child Development 74,* 981–1003.
Brent M (ed.) (1997). *Computational approaches to language acquisition.* Cambridge, MA: MIT Press.
Brown R (1973). *A first language: The early stages.* Cambridge, MA: MIT Press.
Crain S & Thornton R (1998). *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics.* Cambridge, MA: MIT Press.
Darwin C (1877). 'A biographical sketch of an infant.' *Mind 2,* 128–156.
Demuth K (1996). 'Collecting spontaneous production data.' In McDaniel D, McKee C & Cairns H S (eds.) *Methods for assessing children's syntax.* Cambridge, MA: MIT Press. 3–22.
Elman J, Bates E, Johnson M H, Karmiloff-Smith A, Parisi D & Plunkett K (1996). *Rethinking innateness: A connectionist perspective on development.* Cambridge, MA: MIT Press.
Fernald A (1985). 'Four-month-olds infants preference to listen to motherese.' *Infant Behavior and Development 10,* 181–195.

Golinkoff R M K, Hirsh-Pasek K, Cauley K M & Gordon L (1987). 'The eyes have it: Lexical and syntactic comprehension in a new paradigm.' *Journal of Child Language 14,* 23–45.
Hamburger H & Crain S (1987). 'Acquisition of cognitive compiling.' *Cognition 17,* 85–136.
Jusczyk P W (1997). *The discovery of spoken language.* Cambridge, MA: MIT Press.
Lloyd P & Donaldson M (1976). 'On a method of eliciting true/false judgements from young children.' *Journal of Child Language 3,* 411–416.
Lloyd P, Mann S & Peers I (1998). 'The growth of speaker and listener skills from five to eleven years.' *First Language 18,* 81–103.
MacWhinney B (2000). *The CHILDES Project: Tools for analysing talk 1.* Mawah, NJ: Erlbaum.
MacWhinney B & Bates E (eds.) (1989). *The cross-linguistic study of sentence processing.* Cambridge, UK: Cambridge University Press.
Mayer M (1969). *Frog, where are you?* New York: Puffin Books.
McDaniel D, McKee C & Cairns H S (eds.) (1996). *Methods for assessing children's syntax.* Cambridge, MA: MIT Press.
McKee C (1996). 'On-line methods.' In McDaniel D, McKee C & Cairns H S (eds.) *Methods for assessing children's syntax.* Cambridge, MA: MIT Press. 189–208.
Ochs E & Schieffelin B (1995). 'The impact of language socialization on grammatical development.' In Fletcher P & MacWhinney B (eds.) *The handbook of child language.* Cambridge, MA: Blackwell. 73–94.
Rumelhardt D E & McClelland J L (1986). 'On learning the past tense of English verbs.' In McClelland J L, Rumelhardt D E & the PDP Research Group (eds.) *Parallel distributed processing: Explorations in the microstructure of cognition 2: Psychological and biological models.* Cambridge, MA: MIT Press. 216–271.
Shipley E F, Smith C S & Gleitman L R (1969). 'A study in the acquisition of language: Free responses to commands.' *Language 45,* 322–342.
Slobin D I (ed.) (1985). *The cross-linguistic study language acquisition 1–2.* Hillsdale, NJ: Erlbaum.
Slobin D I (ed.) (1992). *The cross-linguistic study language acquisition 3.* Hillsdale, NJ: Erlbaum.
Slobin D I (ed.) (1997). *The cross-linguistic study language acquisition 4–5.* Mawah, NJ: Erlbaum.
Tomasello M & Stahl D (2004). 'Sampling children's spontaneous speech: How much is enough?' *Journal of Child Language 30,* 101–121.
Trueswell J C, Sekerina I, Hill N M & Logrip M L (1999). 'The kindergarten-path effect: Studying on-line sentence processing in young children.' *Cognition 73,* 89–134.
Wulfeck B B (1993). 'A reaction time study of grammaticality judgements in children.' *Journal of Speech and Hearing Research 36,* 1208–1215.

Language Adaptation and Modernization

S Bowerman

This article is reproduced from the Concise Encyclopedia of sociolinguistics pp. 626–630, © 1994, Elsevier Ltd.

Introduction

'Language adaptation' refers to linguistic change, whether deliberate or spontaneous, whose goal is to adapt a language to the changed or changing needs of its speech community. The term is nearly synonymous with 'language modernization,' as the needs of a speech community are often a function of the times. Coulmas (1989: 1) argued that the term 'modernization' is less suitable, as it is often taken to refer to the present day. Language adaptation is a relatively new term in sociolinguistics, but the phenomenon affects all languages at all times.

The field of language adaptation and modernization largely intersects with that of language planning (*see* **Language Planning and Policy: Models**) in that a language is sometimes deliberately adapted to fulfill new roles. Natural language adaptation is spontaneous and is unconsciously undertaken by its own speech community. Intervention is necessary only when a language, previously excluded for some reason from certain domains of use, suddenly needs to be used for those functions. It is therefore desirable to distinguish between linguistic change that comes about naturally and that which is forced.

Factors Influencing (Deliberate) Language Adaptation

Dominance and Domain Exclusion

Many of the world's language communities are dominated in some way by another language community that is politically, economically, or numerically more powerful (cf. Tollefson, 1991: 14). The dominant community's language becomes the medium for the higher domains of language use (*see* **Pragmatics: Linguistic Imperialism**; **Power and Pragmatics**), such as government, media, education, technology, and commerce, to the exclusion of the dominated language.

Golinkoff R M K, Hirsh-Pasek K, Cauley K M & Gordon L (1987). 'The eyes have it: Lexical and syntactic comprehension in a new paradigm.' *Journal of Child Language 14,* 23–45.
Hamburger H & Crain S (1987). 'Acquisition of cognitive compiling.' *Cognition 17,* 85–136.
Jusczyk P W (1997). *The discovery of spoken language.* Cambridge, MA: MIT Press.
Lloyd P & Donaldson M (1976). 'On a method of eliciting true/false judgements from young children.' *Journal of Child Language 3,* 411–416.
Lloyd P, Mann S & Peers I (1998). 'The growth of speaker and listener skills from five to eleven years.' *First Language 18,* 81–103.
MacWhinney B (2000). *The CHILDES Project: Tools for analysing talk 1.* Mawah, NJ: Erlbaum.
MacWhinney B & Bates E (eds.) (1989). *The cross-linguistic study of sentence processing.* Cambridge, UK: Cambridge University Press.
Mayer M (1969). *Frog, where are you?* New York: Puffin Books.
McDaniel D, McKee C & Cairns H S (eds.) (1996). *Methods for assessing children's syntax.* Cambridge, MA: MIT Press.
McKee C (1996). 'On-line methods.' In McDaniel D, McKee C & Cairns H S (eds.) *Methods for assessing children's syntax.* Cambridge, MA: MIT Press. 189–208.
Ochs E & Schieffelin B (1995). 'The impact of language socialization on grammatical development.' In Fletcher P & MacWhinney B (eds.) *The handbook of child language.* Cambridge, MA: Blackwell. 73–94.
Rumelhardt D E & McClelland J L (1986). 'On learning the past tense of English verbs.' In McClelland J L, Rumelhardt D E & the PDP Research Group (eds.) *Parallel distributed processing: Explorations in the microstructure of cognition 2: Psychological and biological models.* Cambridge, MA: MIT Press. 216–271.
Shipley E F, Smith C S & Gleitman L R (1969). 'A study in the acquisition of language: Free responses to commands.' *Language 45,* 322–342.
Slobin D I (ed.) (1985). *The cross-linguistic study language acquisition 1–2.* Hillsdale, NJ: Erlbaum.
Slobin D I (ed.) (1992). *The cross-linguistic study language acquisition 3.* Hillsdale, NJ: Erlbaum.
Slobin D I (ed.) (1997). *The cross-linguistic study language acquisition 4–5.* Mawah, NJ: Erlbaum.
Tomasello M & Stahl D (2004). 'Sampling children's spontaneous speech: How much is enough?' *Journal of Child Language 30,* 101–121.
Trueswell J C, Sekerina I, Hill N M & Logrip M L (1999). 'The kindergarten-path effect: Studying on-line sentence processing in young children.' *Cognition 73,* 89–134.
Wulfeck B B (1993). 'A reaction time study of grammaticality judgements in children.' *Journal of Speech and Hearing Research 36,* 1208–1215.

Language Adaptation and Modernization

S Bowerman

This article is reproduced from the Concise Encyclopedia of sociolinguistics pp. 626–630, © 1994, Elsevier Ltd.

Introduction

'Language adaptation' refers to linguistic change, whether deliberate or spontaneous, whose goal is to adapt a language to the changed or changing needs of its speech community. The term is nearly synonymous with 'language modernization,' as the needs of a speech community are often a function of the times. Coulmas (1989: 1) argued that the term 'modernization' is less suitable, as it is often taken to refer to the present day. Language adaptation is a relatively new term in sociolinguistics, but the phenomenon affects all languages at all times.

The field of language adaptation and modernization largely intersects with that of language planning (*see* **Language Planning and Policy: Models**) in that a language is sometimes deliberately adapted to fulfill new roles. Natural language adaptation is spontaneous and is unconsciously undertaken by its own speech community. Intervention is necessary only when a language, previously excluded for some reason from certain domains of use, suddenly needs to be used for those functions. It is therefore desirable to distinguish between linguistic change that comes about naturally and that which is forced.

Factors Influencing (Deliberate) Language Adaptation

Dominance and Domain Exclusion

Many of the world's language communities are dominated in some way by another language community that is politically, economically, or numerically more powerful (cf. Tollefson, 1991: 14). The dominant community's language becomes the medium for the higher domains of language use (*see* **Pragmatics: Linguistic Imperialism**; **Power and Pragmatics**), such as government, media, education, technology, and commerce, to the exclusion of the dominated language.

Because the dominated languages are excluded from these higher domains, their lexicons, for example, do not keep up with the development of the dominant language as new terms are coined to describe new phenomena and experiences. The need for language adaptation is often recognized when there is a change, or a push for change, in the social structures that keep the dominance relationship between language communities intact. When this happens, and the (previously) dominated language is required as a medium for the higher domains, it is often found to be inadequate in terms of technology, modes of expression, etc.

Domain exclusion also impacts upon a language's speech community on a social level. A language that is excluded from the higher domains of language use is reserved for the lower or more basic domains, i.e., the home and community environment. This means that they will not be perceived as tools for economic advancement and will consequently suffer from low status and prestige (*see* **Language Planning and Policy: Models**; **Pragmatics: Linguistic Imperialism**). The language becomes stigmatized, even in its own community, and members of the language community who seek economic advancement will seek to acquire the dominant language. If the status of a language in such a position is not improved, it may well face obsolescence.

The postcolonial experience of many African and Asian nations in the mid-to-late 20th century was a proving ground for the dominated indigenous languages. The colonial languages had almost invariably been used as media for the higher domains, to the exclusion of the indigenous languages (Phillipson, 1993: 112, Chap. 5). The demise of colonialism meant that many countries were faced with a dilemma: whether to retain the colonial language as a high-domain medium or replace it with one or more indigenous languages. Retaining the colonial language carried unpleasant connotations, but it was already developed enough to cope as the medium of a modern industrial state and was entrenched in that role (Schmied, 1991: 19). The indigenous languages, on the other hand, had been excluded from the higher domains and had neither the status nor the corpus (*see* **Language Planning and Policy: Models**) to cope with their new roles. Although the indigenous languages were often instituted as symbolic official languages, the colonial languages was retained for higher domain use in most cases (Schmied, 1991: 20). Although the dominance of colonial languages, particularly those that offer global access, is all but assured (Laitin, 1992: 181), there is a concerted effort for the functional expansion of indigenous languages.

The Transition from Oracy to Literacy

The move from oracy (*see* **Oracy Education**) to literacy (*see* **Literacy Practices in Sociocultural Perspective**) also creates a need for language adaptation. Indeed, the need for a written language where there was none before reflects a drastic change in the social structure of a community. There are (at least) two ways of fulfilling this need: the adoption of an already written language for these purposes or the adaptation of the oral language to writing. Most societies have had to face this transition at some point in their history, whether this developed naturally or was thrust upon them from outside. Thus, the written language may differ completely from that which is spoken – which, as Coulmas (1989: 13) pointed out, is to be expected as speech and writing serve different communicative functions. Moreover, writing in the initial stages of the transition from oracy to literacy is usually the domain of an elite few. However, the adaptation of an oral language to writing reflects a change in the communicative function of that language, which in modern times has spread to the rest of the society. Whereas the 'undiscovered' world coped for millennia without writing, illiteracy in present times is viewed as a scourge.

Coulmas (1989: 12) noted that literacy does not necessarily stimulate language adaptation *per se*, but serves as a tool for carrying out the process. A written language can easily be codified into a set of rules to be transmitted through education. This in turn affects the acquisition and use (particularly in the higher domains) of a language.

Attitudes and Language Maintenance

As with any language planning measure, it is the attitude of the people for whom language adaptation is to be undertaken that will determine whether or not it is successful. One crucial factor in forming a community's attitude is the degree to which the language will be useful, i.e., the domains to which a particular language gives access (Phillipson, 1993: 79, 80; Cooper, 1989: 100, Chap. 5). Loyalty to a language and culture is another factor affecting the maintenance of a language. Although the revival of Irish Gaelic (Irish) in Ireland has turned out to be little more than symbolic, it is the loyalty of the Irish to their language that led to its extensive adaptation, even though English serves all the language needs of the majority. Languages that are neutral, i.e., do not carry ethnic connotations, are more likely than others to be used as languages of wider communication. The neutrality of Swahili in East Africa led to its being extensively adopted as a national language in Kenya and an official language in Tanzania, and therefore

to the extensive expansion of its functional range (Cooper, 1989: 102).

The Process of Language Adaptation

Two Aspects of Adaptation

The process of equipping a language to deal with the realities of its speech community involves two aspects: one external to the language (involving the speech community) and one internal to the language (involving the structure of the language itself). In order for a language to be successfully adapted, that language must be widely used by its speech community in the domains concerned. This requires not only the elaboration of its lexicon, but also the boosting of its status. Adaptation and modernization through language planning are at an early stage; it is the implementation of language planning measures by a speech community that tells us whether the language has been successfully adapted or not.

Adaptation through Status Planning

A common way of boosting the status of a langauge, and facilitating its adaptation and modernization, is through functional reallocation. Thus, a language previously excluded from the higher domains is allocated to those domains (Cooper, 1989: 99, 118). These measures take time to be implemented: before a language can be properly used in a domain from which it was previously excluded, it must be developed as a medium for that domain.

After their independence from colonial rule in 1962 and 1963, respectively, Tanzania and Kenya wished to establish a national African language in place of the colonial one (Laitin, 1992: 91). Swahili was the natural choice, having few 'tribalist' associations and a long history with the liberation movements (Laitin, 1992: 91–92). In Tanzania (then called Tanganyika), Swahili was proclaimed as the language of state and public functions; shortly thereafter it became the 'major language of Parliament' and was introduced as the 'medium of instruction in all public primary schools' (Massamba, 1989: 63). Although Swahili still faces serious competition from English, its adaptation and modernization have been widely accepted.

Corpus Adaptation and Modernization

Terminology Development In order to fulfill its role as a high-domain medium, a language's corpus (lexicon and grammar) must be sufficiently developed. In cases where a language has been excluded from a domain it now wishes to occupy, its terminology and perhaps modes of expression need to be adapted to its new roles. There are various ways of adapting terminology, taking the target language system into account. It can be achieved through neologism, calquing, transliteration, and meaning extension (Alberts, 1998: 233). Alberts (1998) provided some examples of these four methods of terminology development from South African languages, adapted here in Table 1.

As can be seen from some of the examples in Table 1, terminology development must take the target language into account and we see how calquing and borrowing adapt from the sound system of one language to the other. Grammar is also an important device in language adaptation and, particularly, terminology development. Table 2 shows how new coinages in Irish were applied to existing grammatical structures. O'Baoill (1988: 116) gave the example of the creation of an Irish term for 'lithograph' and some of its derivatives. The new lexeme is a Hibernized borrowing from English, to which productive derivational morphemes are applied.

Languages also reflect their adaptation to a new or changed social order by getting rid of certain elements that tie them to the past. The reform of Turkish under Attatürk in the 1920s saw all Western Farsi (Persian) and Standard Arabic (Arabic) loan words expelled from the standard Turkish lexicon (Cooper,

Table 1 Terminology development in South African languages

Method	*Example*
Calquing: direct translation, especially of compounds	English *cellphone* > Afrikaans *selfoon*
Loan words: borrowing a word, or the sound of a word, by another language	English *cellphone* > Sotho *selula founo*, Zulu *iselula*
	English *computer* > Xhosa *ikompivuta*, Zulu *ikhomputha*
Transliteration: use of the classic stem to coin an equivalent	English *silver* > Xhosa *isilivere*, Zulu *isiliva*
Extension of meaning: use of existing words, but broadening the meaning to include more modern concepts	Zulu *bafana bafana* (reduplication) = 'the boys'; extension of meaning to the soccer team, i.e., young male competitors on the international soccer scene
Neologism: new words created from within the target language	English *unleaded petrol*, Afrikaans *ongelode petrol* (Creation of the verb *lood* 'lead,' past participle *gelode* 'leaded' from Afrikaans noun *lood* 'lead')

Table 2 An example of terminology development in Irish

English term	*English derivational morphemes*	*Standardized Irish borrowing*	*Irish derivational morphemes*
lithograph		*liteagraf*	
	-er		*-ái*
	-y		*-áiocht*

1989: 154, citing Heyd, 1954). Similarly, some sectors of the French government have tried to block the number of loan words entering into French usage (Appel and Muysken, 1993: 164).

Grammatical Development The development of a grammar to suit modern or high-domain purposes often results in a simplification of grammatical forms that are considered unwieldy, anachronistic, and unsuited to scientific and commercial discourse. Grammatical development may mean bringing the grammar of a language more into line with that of a dominant language. For example, there have been some attempts to bring the grammar of Irish more into line with that of English (Greene, 1972: 15). Developing a standard grammar for Hebrew (Modern Hebrew) from Biblical Hebrew processes, as well as innovation, in Israel spurred the creation of a multitude of new word forms. Grammar can especially be a problem in situations where there are a number of varieties of a language. In order for the language to be adapted to modern roles, it requires codification and therefore standardization (*see* **Standardization** (03048)). Thus, where more than one form is found across varieties, one must be chosen as the standard. This applies to all aspects of the corpus.

Orthography and Spelling Orthography and Spelling are two more aspects of the corpus, which are taken into account when conscious adaptation and modernization take place. Orthography refers to the script that is used to represent the spoken language, and spelling refers (roughly) to the way it is used.

When a language is selected for adaptation and modernization, an orthography needs to be selected for it. If the language has never been written before, an orthography must be created – a process known as graphization. If, however, an orthography already exists, a decision must be made as to whether this system should be retained or a new system found. This process is known as regraphization (Cooper, 1989: 125, 154). Christian missionaries graphized most of the oral languages of southern Africa in the 19th and 20th centuries. Generally, a Roman script was imposed and the spelling system of the missionaries' nationalities informed the orthography and spelling. Whereas graphization is one aspect of language adaptation, regraphization is also important. The change in orthography reflects a change in society. An early example of this was the development of the Cyrillic alphabet, so that biblical and ecclesiastical texts could be translated into Slavic languages. More recently, Turkish adopted a Roman script in place of the Persian orthography it had used before (Haugen, 1972: 170, citing Heyd, 1954). Orthography can have powerful symbolic value for a speech community. In the 1930s, the Soviet government imposed the Cyrillic script on non-Russian languages within the USSR (Mesthrie *et al.*, 2000: 391). Since the disintegration of the USSR, however, a number of these languages have had the non-Cyrillic script reinstated. Mesthrie *et al.* (2000: 391) cited the case of Moldavia: in August 1990, the Moldavians held a 'Language Festival,' celebrating the reintroduction of the Latin script. A developing adaptation process is also evident in Mandarin Chinese (Chinese) and Japanese, which are adapting to changing functions by *adding* a Roman orthography, to facilitate the learning of these languages phonetically (Crystal, 1995: 313; Sugito, 1989: 117).

Standardization and Harmonization Standardization and harmonization apply in different ways to the same phenomenon: one language, which has many distinctive varieties in the same area. In some cases, like that of Irish (O'Baoill, 1988: 111) or Swahili (Massamba, 1989: 61), the varieties are widely recognized by academics and speakers as being varieties of one language. In order for the language to be adapted to a function, it needs to be stabilized and codified, and therefore standardized, and an orthographical/terminological/grammatical form is chosen from one of the dialects and held up as the norm. Otherwise, one dialect will be chosen as the standard and all its forms held up as the norm by society as well as language planners (Cooper, 1989: 134).

Harmonization is a form of standardization, but operates across varieties that are regarded as closely related, but distinct, languages. In essence, harmonization is an attempt to build one standard language from a variety of closely related languages (and all of their respective varieties). The harmonization (also 'unification') question in South Africa was revisited in 1989 by Neville Alexander (1989: 64), who proposed the harmonization of the seven major indigenous languages into two standards:

1. Standard Nguni, comprising Xhosa, Zulu, Ndebele, and Swati;
2. Standard Sotho, comprising Southern Sotho, Northern Sotho (Pedi), and Tswana.

Table 3 Equivalent terms in Xhosa, English, and Zulu

Xhosa	*Zulu*	*English*
ifoni	*ithelefoni*	telephone
imoto	*imoto*	motor car
umbane	*ugesi*	electricity
ilayibri	*ilabhulali*	library
ikompiyuta	*ikomputha*	computer

Source: Adapted from Reader's Digest (1985) *South African multi-language dictionary*, with permission.

Harmonizing the indigenous languages of South Africa would eliminate the need to adapt and modernize every one of the seven languages that make up Sotho and Nguni. The harmonization of these languages could begin with new terminology, which can be coined as illustrated above. Table 3 illustrates similarities and differences between the Nguni languages.

The similarity between these two languages is striking and shows the close relation of the languages to each other. Differences in vocabulary and pronunciation appear minor, but are significant to their speech communities, which may be concerned that linguistic assimilation means political assimilation (Msimang, 1998: 167). Those involved in language harmonization will have to liase extensively across languages to avoid being perceived to value one variety or language group above another.

Conclusion

Language adaptation and modernization are essential to the evolution and, indeed, survival of any language. The linguistic change brought about through adaptation reflects the dynamics of the society in which it is used. The fact that intervention is sometimes necessary links the field with that of language maintenance: the adaptation of a language to fulfill modern functions is surely a mechanism for language maintenance. It ensures that a language remains valued by its own community and others.

See also: Language Development: Overview; Language Planning and Policy: Models; Literacy Practices in Sociocultural Perspective; Oracy Education; Power and Pragmatics; Pragmatics: Linguistic Imperialism; Standardization.

Bibliography

Alberts M (1998). 'The harmonisation of Nguni and Sotho.' In Prah K K (ed.) *Between distinction and extinction: The harmonisation and standardisation of African anguages*. Johannesburg, South Africa: Witwatersrand University Press.

Alexander N (1989). *Language policy and national unity in South Africa, Azania*. Cape Town, South Africa: Buchu Books.

Appel R & Muysken P (1993). *Language contact and bilingualism*. London: Arnold.

Cooper R (1989). *Language and social change*. Cambridge, UK: Cambridge University Press.

Coulmas F (1989). 'Language adaptation.' In Coulmas F (ed.) *Language adaptation*. Cambridge, UK: Cambridge University Press.

Crystal D (1995). *The Cambridge encyclopedia of language*. Cambridge, UK: Cambridge University Press.

Greene D (1972). *The Irish language*. Cork, Ireland: Cultural Relations Committee of Ireland.

Haugen E (1972). *The ecology of language*. Stanford, CA: Stanford University Press.

Heyd U (1954). *Language reform in modern turkey*. Jerusalem: Israel Oriental Society.

Laitin D (1992). *Language repertoires and state construction in Africa*. Cambridge, UK: Cambridge University Press.

Massamba D (1989). 'An assessment of the development and modernization of the Kiswhahili language in Tanzania.' In Coulmas F (ed.) *Language adaptation*. Cambridge, UK: Cambridge University Press.

Mesthrie R, Swann J, Deumert A & Leap W (2000). *Introducing sociolinguistics*. Edinburgh, UK: Edinburgh University Press.

Msimang T (1998). 'The nature and history of harmonisation of South African languages.' In Prah K K (ed.) *Between distinction and extinction: The harmonization and standardization of African languages*. Johannesburg, South Africa: Witwatersrand University Press.

O'Baoill D (1988). 'Language planning in Ireland: The standardization of Irish.' *International Journal of the Sociology of Language 70*, 29–52.

Phillipson R (1993). *Linguistic imperialism*. Oxford, UK: Oxford University Press.

Reader's Digest (1985). *South African multi-language dictionary.* Cape Town, South Africa: Reader's Digest.

Schmied J (1991). *English in Africa: An introduction*. London: Longman Linguistics Library.

Sugito S (1989). 'Lexical modernization of Japanese.' In Coulmas F (ed.) *Language adaptation*. Cambridge, UK: Cambridge University Press. 116–126.

Tollefson J (1991). *Planning language, planning in equality: Language policy in the community.* Essex, UK: Longman.

Language and Dialect: Linguistic Varieties

P Trudgill, Universite de Fribourg, Fribourg, Switzerland

Max Weinreich's (unpublished) dictum that "a language is a dialect with an army and a navy" contains much truth, but it is a partial truth. Heinz Kloss (1967) agrees that the notion of 'a language' is as much a political, cultural, and historical notion as it is a purely linguistic notion, but he also points out that languages are of two main sociolinguistic types. These types are called in German *Abstandsprachen* and *Ausbausprachen*. Abstand languages can be called in English 'languages by distance.' They are linguistic varieties that are regarded as separate language by reason of their linguistic distance from all other languages. Basque is an isolate with no known relatives and is extremely distant in terms of its linguistic structures from all other languages, including its geographical neighbors. The status of Basque as a language is thus assured on purely linguistic grounds.

Ausbau languages are in English called 'languages by extention, or construction.' They are linguistic varieties that are regarded as distinct languages for political, cultural, social, and historical as well as linguistic reasons. Ausbau languages depend on cultural factors for their status because they are associated with geographical dialect continua. Polish, Slovak, and Czech are Ausbau languages which together form the West Slavic dialect continuum. They are mutually intelligible, especially Polish and Slovak, and Slovak and Czech, but they form three separate languages because there are three separate standardized norms with their own agreed-upon and accepted orthographies, shared histories, traditional literatures, widely used dictionaries, and grammar books – and their own nation states. Einar Haugen (1966) argues that because of these separate orthographies, literatures, and histories, each of these three standard varieties has the characteristic of *autonomy* – they are independent varieties with, as it were, an existence of their own.

This provides a definition of an Ausbau language as: an autonomous standardized variety together with all the nonstandard dialects from that part of the dialect continuum that are heteronomous with respect to it, i.e., dependent on it. With the term *heteronomy*, or 'dependence,' Haugen refers to facts such as that a Slovak and a Pole from opposite sides of the national frontier may be able to converse with one another readily, but one will go home and read Slovak newspapers in Slovak, while the other will go home and watch Polish television.

Autonomy and heteronomy are cultural and political phenomena, which entails that Ausbau languages are cultural and political constructs. Ausbau languages are therefore potentially temporary entities. They may become dialects, through loss of autonomy, as in the case of Scots, which was formerly a distinct language but is now widely though not universally regarded as a dialect of English. And dialects can become languages, by gaining autonomy, such as when Norwegian lost its heteronomy with respect to Danish, and Macedonian its heteronomy with respect to Bulgarian and Serbian.

See also: Accent; Dialect Chains; Endangered Languages; English: World Englishes; Haugen, Einar (1906–1994); Identity and Language; Linguistic Ethnonationalism; Nationalism and Linguistics; Politics and Language: Overview; Researching Naturally Occurring Speech; Sociolect/Social Class; Standard and Dialect Vocabulary; Subjective and Objective Measures of Variation.

Bibliography

Haugen E (1966). 'Dialect, language, nation.' *American Anthropologist 68,* 922–935.

Kloss H (1967). 'Abstand languages and Ausbau languages.' *Anthropological Linguistics 9,* 29–41.

Language as an Object of Study

V Žegarac, University of Luton, Luton, UK

> Words are the leaves of the tree of language, of which, if some fall away, a new succession takes their place.
> – Field Marshall John French

What are the properties of language that make it similar to a tree? Could language exist without words in the way a tree may exist without leaves? Could words exist without the tree of language? One way to address these and other similar questions might be to make a list of the main properties of trees. The next step might be to look systematically for identical or similar properties in human languages. This would mean that our knowledge of botany would be guiding our scientific investigation of language. As fundamental differences between trees and languages are many and easily observable, it is evident that this research program does not hold enough promise to be worth pursuing. However, this general approach illustrates an important point about the methodology of scientific research in general and linguistic investigation of language in particular: research into one poorly understood object of study proceeds on the assumption that it shares its essential properties with some other (type of) object which is already better understood. In fact, some major landmarks in the development of modern linguistics are based on four conceptions of language as an object of study: language as a social fact, language as behavior, language as a mental organ, and language as an abstract object.

Language as a Social Fact

The view of language as a social fact, as suggested by Ferdinand de Saussure (1916), marked a shift away from the predominantly historical interest in language toward the study of language as a system at the current stage of its development. Saussure believed that important insights into language could be gained if it was investigated from the point of view of its users, who typically do not know anything about its historical development. Since language use reflects the structure of language as a system of units at a particular stage of its development, language should be studied from the synchronic, rather than the diachronic, perspective. Saussure compared a given stage in the development of a language with the configuration of pieces on the chess board at a given stage of a game of chess. Just as the value of each piece is determined by its position relative to all the other pieces on the board (at a given stage of the game), a particular language at a given stage of development is a system of language units, or language signs (Saussure's term is *signe linguistique*), whose values are determined by their position in relation to all the other signs in the language system at that particular stage. And just as any change in the position of a given chess piece alters the values of all the other pieces and affects the system as a whole, the change that directly affects one language sign (potentially) affects the language system as a whole. For example, although the Serbian word 'jeftin' has the same meaning as the English word 'cheap,' it does not have the same value in the language system, because, in English, 'cheap' contrasts with 'inexpensive,' whereas in Serbian the same contrast is not lexicalized, so that it is either ignored, with 'jeftin' covering both meanings, or it is conveyed in some more elaborate way. The value of the Serbian word 'jeftin,' – its position in the language system – would change significantly if the equivalent of the English 'inexpensive' were to become lexicalized in Serbian.

In Saussure's view, language is a social fact akin to other social institutions, such as the legal system or the dress code. Although people's use of language reflects its systematic character, the structure of language as a system is not directly observable or consciously represented by its users. Saussure posited a distinction between observable linguistic behavior, what he called *parole* (speaking), and the underlying language system, what he called *langue* (language). In his view, the main task of linguistics was to discover *langue* – the language system – by investigating *parole* – language in use. Saussure's view of languages as a social fact is consistent with Émile Durkheim's sociological theory (popular at the time when Saussure was developing his ideas on language), in which social facts are characterized as representations in the 'collective mind,' rather than as material or psychological entities.

Saussure's concept of *langue* as a social fact, and as the proper object of linguistic investigation, had an unfortunate consequence. If *langue* is a system of relations between language signs, which are both external to its users and not directly observable in the manifestations of *langue* in use (because they are part of the 'collective mind'), there is no place for the category of sentence in linguistic analysis. In other words, it follows from Saussure's position that sentences are an aspect of *parole* rather than *langue*. Therefore, they are not a part of the systematic character of a language. So, although we readily produce and comprehend utterances of sentences that we have never produced or heard before and, although

we assign them meanings in a systematic way, Saussure's view of language as a social fact did not adequately accommodate this important aspect of language.

Language as Behavior

The view of language as behavior emerged in American descriptivist linguistics under the influence of positivism (in philosophy) and behaviorism (in psychology). An important methodological stance of logical positivism is that scientific research must be based either on statements that are tautologically true, such as: either P or not P, or they must be based on direct observation. The immediate consequence of this view for the scientific study of language was that all analyses and generalizations should be based strictly on the description of observable linguistic behavior. The main methodological stance of behaviorist psychology was that scientific theories could be refuted only by observable phenomena. Hence the descriptivists' view that insights into language should be firmly grounded in observable data – the products of linguistic behavior – and that they should not be informed by the researcher's introspection or the native speaker's intuition. This meant that more attention was paid to particular languages than to the universal properties of human language, as we can only have direct data from particular languages. However, although descriptivist linguists initially thought of human languages as infinitely diverse, by the 1950s their interest had shifted towards the possibility of devising a discovery procedure, a mechanism that would take a corpus of utterances as input and yield a grammar as output (Harris, 1951).

The most important theoretical tenet of behaviorism was that all human behavior could be explained in terms of stimuli and responses to stimuli without reference to mental structures and processes. This stance had two major implications for the study of language: first, language was characterized as "the totality of utterances that can be made in a speech community" (Bloomfield, 1928). Second, since meaning is not directly observable and it could not be studied without reference to intuitive, introspective judgments, it was considered not to be the proper object of linguistics. Though to a lesser extent than Saussure's conception of language as a social fact, the view of language as the product of behavior (especially in the early stages) also tended to focus on the phonological and morphological analysis of language, rather than on the less directly observable structural relations between words within sentences.

Language as a Mental Organ

The view of language as a mental organ is central to the most influential modern linguistic theory, the generative approach to language, developed by Noam Chomsky since the 1950s and his many followers (see Chomsky, 1986, 1988, 2000). Although the system of ideas that characterizes present-day generative linguistics is very complex, the basic assumptions on the nature of language are clearly established and they appear to be fundamentally commonsensical.

The idea that language is a mental organ is best illustrated by an analogy with other biologically specified properties of human beings. The development of any organism is partly determined by its genetic makeup and partly by environmental factors. For example, the onset of puberty and the height of a person are genetically determined, but they are also affected to some limited extent by external factors, such as nutrition. Given the diversity of environmental circumstances in which people grow up, and the uniformity in their physical appearance, there can be hardly any doubt as to the primacy of genetic makeup over the environment. The same observation carries over to our mental capacities. For example, numerical ability is unique to humans, who develop it in a uniform way regardless of the differences in their sociocultural and physical environments. Similarly, the design features of language, such as recursion, arbitrariness, and duality of patterning, cannot be explained as the result of environmental factors. Therefore, we should assume that both our ability to count and our linguistic ability are genetically specified. In other words, just as arms and legs are physical organs, language is a mental organ. The interaction with the environment triggers the development of language, but the environmental inputs are simply not sufficiently structured for it to be plausible to assume that language development involves the construction of more and more detailed mirror images of the external world. Other comparable capacities would include our musical ability and the ability to construct scientific theories.

If language is a mental organ, then linguistics should be seen as a discipline of psychology, ultimately biology. It should aim to discover the abstract properties of language that make it possible to answer the following questions:

1. What do we know when we know language?
2. How is knowledge of language acquired?
3. How is this knowledge of language put to use?

In this view, the knowledge of language is characterized as a (generative) grammar: a finite system of interacting rules and principles that specify (more

technically, generate) an infinite number of expressions, each of which is a structural configuration (e.g., of words into phrases and sentences) associated with a phonetic form and meaning. At the core of our ability to acquire language lie some universal innate properties of grammar, which guide the child's language development. The innate properties of grammar constrain the range of possible grammars of language and enable the child to find out which of the possible grammars is best evidenced by the linguistic behavior of those around him. This genetic endowment enables the child to go beyond the evidence presented by the data in determining the grammar. For instance, a child acquiring English figures out that 'John is too stubborn to talk to' means, roughly, 'John is so stubborn that nobody can talk to him' without explicit instruction and without being distracted by the possible analogy with utterances like 'John is too stubborn to talk to Bill,' which would be misleading. Therefore, evidence from language acquisition supports the view that the child constructs the grammar on the basis of innate mental structures that heavily constrain the range of available choices at any given stage of his linguistic development. The question of how knowledge of language is put to use is investigated by designing 'performance models' based on the mental grammar.

The research program of Chomskyan generative grammar differs in two important respects from those associated with the view of language as a social fact and the view of language as behavior. First, there is a shift of emphasis from the study of individual languages to the study of Universal Grammar (UG), where the latter is another name for language as a mental organ (in terms of the properties it identifies as the set of grammars that a human being can learn under the conditions in which language acquisition normally takes place). Second, language as an object of study is an element of the mind, which Chomsky calls Internalized language (I-language). The approaches that consider language to be a social fact or behavior and the products of behavior are concerned with the study of language as external to the minds of its speakers. In Chomsky's terms, they are concerned with the study of Externalized language (E-language). Chomsky argues that E-languages (which correspond closely to what people normally mean when they speak about English, French, German, Chinese, Swahili, etc.) are epiphenomenal. In other words, they are complex constructs resulting from the interaction of various sociopolitical and historical factors. Therefore, understood in this way, a particular language like English or French is not a proper object of scientific investigation. Although people ordinarily speak about particular languages as individual objects that exist independently of individual speakers (in utterances like 'French is the language of diplomacy.'), Chomsky insists that this shift of focus from E-language to I-language is a shift towards the commonsense view of language, because when we say that a person knows a language we usually mean that they know how to relate sound patterns with meanings in a particular systematic way.

Language as an Abstract Object

In contrast to Chomsky's mentalist conception of language, Katz (1981) presents a detailed argument to support the view of language as an abstract object. The distinctive characteristics of abstract objects are (a) that their existence is independent of the existence of the mind, and (b) that they do not occupy a position in space and time. Presumably, natural numbers, which underlie the laws of nature, would exist even if human or other intelligence did not, and they are not located at particular places and times. In the light of these observations a natural language, such as English or Japanese, might seem a poor candidate for the category of abstract object. However, Katz (1981) argues that the task of writing the grammar of a language is too stringently constrained on Chomsky's requirement "that a grammar represent a language only in the particular form that knowledge of the language takes when such knowledge is realized in the human mind or brain" (Katz, 1981: 92). Arguably, this requirement has some unwelcome implications for linguistic analysis. For example, let us say that two grammars of a given language (G1 and G2) have been written. G1 reflects the speaker-hearer's knowledge better than G2, but G2 is simpler and more elegant than G1. G1 and G2 are equivalent in all other respects. The question is: 'Is G1 to be preferred to G2, or is G2 to be preferred to G1?' In Katz's view it would be a mistake to reject G2 just because it is less psychologically plausible than G1. He points out the importance of the distinction between the knowledge of something and the object of knowledge itself. For example, natural numbers are distinct from the different ways in which they are represented in various calculators. By the same token, the speaker's knowledge of language is distinct from language itself. Chomsky (1986: 49) rejects this view with the following analogy between language and the heart as biological organs: the heart may be simply and elegantly described as a pump, disregarding what actually makes it beat as a part of the organism, but it does not follow that such a description would be superior to one which looks at the actual anatomy and physiology of the human heart.

Therefore, if G2 is simpler and more elegant than G1, but G1 is more plausible as a psychologically real description of the speaker-hearer's knowledge of their language, then it is obvious that G2 should not be preferred to G1.

Katz also challenges Chomsky's view that individual languages are not proper objects of scientific study because they are sociopolitical constructs. In his view, this is "like claiming that the concept of number is not an object of mathematics but a sociopolitical one." (1981: 79–80). Chomsky (1986: 47) rejects this objection out of hand by describing it as a "curious conclusion" (for other criticisms and a defense against them, see Katz, 1985). Although some criticisms of Katz's view of language seem very convincing, a reasonably strong case can be made in support of the conception of language as an abstract object. On the one hand, the idea that language can be seen and studied scientifically as an abstract object does not entail that it is not also a psychological object. Hence, it is far from obvious that the only interesting and useful scientific approach to the study of language is psychological (biological). In other words, the conceptions of language as a mental organ and language as an abstract object might both provide the basis for credible scientific research. Moreover, the scientific study of various physical organs, such as the human heart, can easily draw on readily available evidence of their structure. By comparison, the evidence for the analysis and description of language as a mental organ is rather scant. Thus, if an organ, e.g., the human heart, were not available for relatively direct observation, then it would be perfectly sensible to investigate it by trying to figure out the best abstract model in the hope that it would also turn out to be the one that corresponds most closely to the real thing. In fact, much work within Chomsky's generative approach in linguistics seems to proceed in this way. On the other hand, Katz (1985: 199–200) argues that theories of language grounded in individual psychology are not abstract enough to provide plausible explanations of language structure. His discussion of analytic sentences is particularly convincing. Analytic sentences, such as 'Nightmares are dreams.' and 'Flawed gems are imperfect.' are necessarily true in virtue of the semantic level of grammars. The Chomskyan approach, which characterizes the grammar in psychological terms, cannot provide an adequate account of necessary truths. 'Nightmares are dreams.' and 'Flawed gems are imperfect.' are true sentences, regardless of our psychological (biological) makeup, but, in the Chomskyan approach, these sentences have to be analyzed as necessarily true, because humans are psychologically (biologically) designed so that we cannot think of them as anything other than true statements. While this and related issues remain open for debate, there can be no doubt that Chomsky's psychological approach to language remains the most influential perspective from which it is studied within linguistics.

Bibliography

Bloomfield L (1928). 'A set of postulates for the science of language.' *Language* 2. In Joos M (ed.) (1957) *Readings in linguistics*. Washington, DC: American Council of Learned Sciences, 1957.

Bloomfield L (1933). *Language*. New York: Holt.

Chomsky N (1986). *Knowledge of language: its nature, origin and use*. New York: Praeger.

Chomsky N (1988). *Language and problems of knowledge: the Managua lectures*. Cambridge, MA: MIT Press.

Chomsky N (2000). *New horizons in the study of language and mind*. Cambridge: Cambridge University Press.

de Boysson-Bardies B (1999). *How language comes to children*. Cambridge, MA: MIT Press.

de Saussure F (1916). *Cours de linguistique générale*. Bally C & Sechaye A (eds.). Paris: Payot, 1968.

Harris Z S (1951). *Structural linguistics*. Chicago: The University of Chicago Press.

Katz J (1981). *Language and other abstract objects*. Oxford: Basil Blackwell.

Katz J (1985). 'An outline of Platonist grammar.' In Katz J (ed.) *The philosophy of linguistics*. Oxford: Oxford University Press.

Pinker S (1994). *The language instinct: the new science of language and mind*. London: Penguin Press.

Sampson G (1980). *Schools of linguistics: competition and evolution*. London: Hutchinson.

Language Attitudes

J Edwards, St Francis Xavier University, Antigonish, NS, Canada

> One grievous failing of Elizabeth's was her occasional pretty and picturesque use of dialect words – those terrible marks of the beast to the truly genteel.
> Thomas Hardy (*The mayor of Casterbridge*, 1886)

> It is impossible for an Englishman to open his mouth without making some other Englishman hate or despise him.
> George Bernard Shaw (*Pygmalion*, 1912)

> The operative class, whose massacre of their mother tongue, however inhuman, could excite no astonishment.
> Thomas Hamilton (*Men and manners in America*, 1833)

Introduction

The most basic theme in modern social psychology is the importance of perception. We do not react to the world on the basis of sensory input but, rather, in terms of what we perceive that input to mean. This is the foundation of all our social constructions, of all our individual and group relationships, and it is a foundation that reflects – in ongoing fashion – our accumulated social knowledge. Perception is the filter through which sensory data are strained, and the establishment and maintenance of this filter is culturally specific and – within social groupings – individualized to a greater or lesser extent. For example, because every individual has accumulated a unique set of experiences, each set of perceptual spectacles is itself special to some extent. At the same time, there are many social perceptions that group members hold in common: at one level, we can think of these as stereotypes, at another as culture itself. At a group level, the most important feature has to do with the acceptance of shared understanding. In a great many instances, this is the only sort of understanding to which we have access, and we sometimes fall prey to the idea that **our** understanding is, in fact, the only sort possible. Nonetheless, we are aware (occasionally, at least) of the relative nature of social knowledge when we realize that others do not sort out their world in quite the same way as we do.

Perceptions operate in societies, most of which are stratified in various ways and in which – to put things bluntly – power and status are often able to translate social difference into social deficiency. Because language is one of the traditionally important social markers, it is not surprising that the study of attitudes has a central position in the social psychology of language. Nor is it surprising that the results of this study shed some light on the generalities just mentioned. It is easy, for example, to demonstrate that great variation exists in terms of people's reactions to (or evaluations of) different accents and dialects. The question, of course, is what to make of this variation.

The Basis of Judgment

Social preferences and prejudices concerning language varieties are long-standing and of continuing potency. This is because views of language correspond to views of the social status of language users; in this sense, language (or dialect or accent) provides simple labels which evoke social stereotypes that go far beyond language itself.

We can begin here by observing that even dictionary definitions of dialect and accent help to sustain the view that nonstandard language is less correct language. The *Oxford English Dictionary* (*OED*), for example, notes that dialect may be considered as "one of the subordinate forms or varieties of a language arising from local peculiarities of vocabulary, pronunciation, and idiom." This definition is implicitly held by those for whom the term 'dialect' conjures up an image of some rustic, regional speech pattern; for such people, the 'standard' variety (see below) is not really a dialect at all; rather, they consider it as the 'correct' form from which all others diverge. On accent, the *OED* is not much better; it is, we find, a mode of utterance which 'consists mainly in a prevailing quality of tone, or in a peculiar alteration of pitch, but may include mispronunciation of vowels or consonants, misplacing of stress, and misinflection of a sentence.'

But what is the basis for language attitudes and judgments? Three broad possibilities suggest themselves.

Intrinsic Difference

One possibility is that language attitudes reflect intrinsic differences across and within language varieties themselves. That is, the reason for variety A being evaluated so much more favorably than variety B is simply that A is a linguistically superior form. Although this is a view that has had considerable historical support, and while it remains common at a popular level in virtually all linguistically stratified societies, linguists have convincingly demonstrated that to see languages or dialects in terms of innate superiority or inferiority is to profoundly misunderstand the nature of human language itself:

> just as there is no linguistic reason for arguing that Gaelic is superior to Chinese, so no English dialect can be claimed to be linguistically superior or inferior to any other . . . There is no linguistic evidence whatsoever for suggesting that one dialect is more 'expressive' or 'logical' than any other, or for postulating that there are any 'primitive', 'inadequate' or 'debased' English dialects. (Trudgill, 1975: 26)

A very good demonstration of this was provided by William Labov (Labov, 1976). He studied Black English (in the United States), which makes an excellent test case for dialect validity since it had for so long been rejected by the white middle class, and since its speakers were victims of a prejudice that went well beyond language alone. If it could be shown that Black English was not some debased variety, this would go some way towards establishing linguistic integrity for all dialect varieties.

There were three strands to Labov's work. First, he justly criticized earlier studies which had elicited Black English from youngsters through interview techniques which were both unfamiliar and intimidating; these were hardly likely, he noted, to produce normal, conversational samples. Second, Labov reminded us of what casual observers had known for a very long time – the Black community is verbally rich and, like other 'oral' cultures worldwide, supports and rewards those who are particularly linguistically gifted. Third, and most important of all, Labov demonstrated the regular, rule-governed nature of Black English. Rules are of course an essential feature of language, and if it were possible to demonstrate formal grammatical regulation within Black English – not necessarily identical, of course, to regulation in other dialects – then charges of inaccuracy and sloppiness would become groundless, and it could not be dismissed as some 'approximation' to 'proper' English.

As an example, here is one of the rules Labov described for Black English, a practice called 'copula deletion.' In a sentence such as *she the first one started us off*, the *is* is not present. Does this mean that the Black speaker who uttered it is unaware of this verb form? Well, consider another sentence taken from a Black corpus: *I was small then* – here, the past tense of the verb 'to be' appears. Why should it figure in this sentence but not in the former one? Labov's work revealed the regularity governing this linguistic behavior: in contexts in which the standard variety permits verbal contraction (e.g., *they are going* can become *they're going*), Black English allows deletion (e.g., *they are going* can become *they going*). We have, therefore, a rule – slightly different from the one obtaining in the standard, but no less logical. The regularity is further evidenced by the fact that, in contexts in which the standard does not allow contraction, there is no Black English deletion: *he's as nice as he say he's* is patently incorrect according to the rules of the standard variety, which outlaws contraction in the final position. In like fashion, it is incorrect to say *he's as nice as he says he* in Black English.

The import of this sort of work is clear: there are no substandard language varieties. There are standard dialects in many languages, and so it logically follows that all others must be nonstandard – but this latter term is not pejorative in any technical linguistic sense. A standard dialect is, roughly, that spoken by educated people, and it is the form usually found in writing. Its position derives from history and from the social standing of its speakers. If York instead of London had become the center of English power, for instance, then the traditional BBC newsreader would sound quite different, and schoolteachers would be promoting another form of 'correct' English in the classroom.

Aesthetic Difference

Another possibility might be that language varieties – although not to be seen (grammatically or 'logically') in terms of 'better' or 'worse' – do actually vary in their aesthetic qualities. Perhaps, then, more favorable attitudes might attach to those varieties that sound better, or more mellifluous, or more musical. Many years ago, for instance, standard English was defended as 'one of the most subtle and most beautiful of all expressions of the human spirit' (Chapman, 1932: 562) and, a little later, Henry Wyld, a linguist at the University of Liverpool, wrote:

> If it were possible to compare systematically every vowel sound in RS [Received Standard English – i.e., what we now more usually call RP, Received Pronunciation] with the corresponding sound in a number of provincial and other dialects, assuming that the comparison could be made, as is only fair, between speakers who possessed equal qualities of voice, and the knowledge how to use it, I believe no unbiased listener would hesitate in preferring RS as the most pleasing and sonorous form, and the best suited to be the medium of poetry and oratory. (Wyld, 1934: 610)

I need hardly say that such sentiments were (and are still, of course) not restricted to those speaking in and for English. Can we put to the test, however, the belief that dialect A is more aesthetically pleasing than dialect B?

Recent studies have compared an 'inherent value' hypothesis here with an 'imposed norm' one. The former term suggests, as Wyld did, that aesthetic qualities are intrinsic, while the latter holds that

they are attached or imposed by the listener, who, in hearing a standard (for instance), considers it mellifluous because of the status of its speakers. In one investigation, Welsh adults listened to European French, educated Canadian French and working-class Canadian French voices. Asked to rate the pleasantness and prestige of the voices, the judges – who were virtually ignorant of French – did not single out any of the three varieties. Earlier studies had shown, however, a clear aesthetic preference among French speakers for European French. In another experiment, British undergraduates who knew no Greek evaluated the aesthetic quality of two Greek dialects, the Athenian and the Cretan: the former is the prestige standard form, while the latter is a nonstandard variant of generally low status. As in the French study, no significant differences between the two dialects were found. If anything, there was a slight tendency for these British students to rate the Cretan variety as more pleasant and prestigious than the Athenian (Giles *et al.*, 1974, 1979).

The important element in these demonstrations is that the judges were unaware of the social connotations possessed by the different varieties in their own speech communities. The implication is that, if one removes (experimentally) the social stereotypes usually associated with given varieties, aesthetic judgments will not be made which favor the high-status standards. Anyone who watches a film or a play in which (for example) a woman dressed as a duchess speaks with a Cockney accent can appreciate the point here: someone in the audience who had an understanding of English, but not of more subtle intralinguistic variation and convention, would miss a great deal of the comedic effect. The norms here are imposed by those in the know, and the stereotypes which link beauty, or harshness, or comedy to a particular set of sounds are unavailable to others. (None of this, I hasten to add, rules out purely individual preferences: I may think Italian is the most beautiful language, you may believe that Gaelic is unrivalled – but we can agree to differ on a matter of subjectivity.)

Social Perception

We now turn to a third possibility, which – because we have eliminated the first two – becomes the most likely: the variant evaluations found in the social laboratory and on the street reflect perceptions of the **speakers** of given varieties; any qualities – 'logical' or aesthetic – of the varieties themselves are, at best, of very secondary importance. Thus, listening to a given variety acts as a trigger or a stimulus that evokes attitudes (or prejudices, or stereotypes) about the community to which the speaker is thought to belong.

A general moral here seems to be this: when social stratification is associated with linguistic variation, the variety used by those with social clout will commonly be perceived as (grammatically, lexically, or phonologically) superior to nonstandard forms. There are two exceptions to this rule, and they occur at opposite ends of the status continuum. First, extremely high-status varieties may seem affected and generally over the top. Second, there is a 'covert prestige' possessed by working-class speech, with its positive associations of masculinity. Research has found that both working-class and middle-class males claimed to use nonstandard forms even when they did not customarily do so (Edwards, 1989; Trudgill, 1983). The actual use of such forms, by generally standard-speaking individuals, is most likely when the speaker wants to appear forceful, direct, and unambiguous. An example: a friend of mine, a middle-aged, upper-middle-class American university professor (male), was being pressed by colleagues (also male) on a departmental matter. Finally, he ended a rambling and inconclusive discussion – one that had unfolded according to the norms of standard, educated American English – by smiling broadly and saying, "Listen, boys, you **know** the' ain't no way I can do it." All-male social gatherings, as at least half the readers will know, often produce such examples.

Variation in Social Attitudes

If we reject linguistic and aesthetic arguments as the real underpinnings of variations in language (and speaker) evaluation, we have only cleared away some annoying underbrush. It is very useful to know that the basis for language evaluation rests upon social convention, but we must make more detailed inquiries about the assessment of language varieties. We come then, to language attitudes *per se*.

While many regional or class dialects are perceived unfavorably, there is more than a simple dichotomy between good (standard) and bad (nonstandard). Urban speech patterns (such as, in Britain, those of Birmingham, Glasgow, and Liverpool) are generally seen as more unpleasant than are rural varieties (Devon English, for instance). Again, the basic explanation rests upon social connotation: for many, rural areas have a charm that is absent in heavily urbanized, industrialized centers. So rural dialects, although not as prestigious as the standard, may be viewed more favorably than urban varieties (which are often tied more closely to class differences, as well) (Wilkinson, 1965; Trudgill, 1975).

Wyld (above) hinted at an important problem in the comparison of language varieties. If I wish to find out which of two dialects is the more pleasant (for example), and if I therefore record a speaker of each and have the voices judged by listeners, to what can I reasonably attribute any differential ratings that may be found? Are they due to features of the dialects themselves? Might they not also be reactions to purely individual qualities of voice – tone, pitch, rhythm, pace, and so on? A way around this difficulty was devised in the 1960s by Wallace Lambert. In his 'matched guise' method, judges evaluate a tape-recorded speaker's personality after hearing him or her read the same passage in each of two or more languages, dialects, or accents. The fact that the speaker is, for all 'guises,' the same person is not revealed to the judges (who typically do not guess this). Any variations in the judges' ratings can then be considered as reflections of their stereotyped reactions to the different language varieties, since potentially confounding individual variables are of course constant across guises (see Lambert *et al.*, 1960; Edwards, 1989).

A large number of studies have shown that evaluations of different varieties vary in nonrandom ways. Thus, standard accents and dialects usually connote high status and competence; regional, 'ethnic,' and lower-class varieties are typically associated with greater speaker integrity and attractiveness. Ellen Bouchard Ryan and her colleagues (Ryan *et al.*, 1982, 1984) have attempted to summarize the many language attitude studies by providing an "organizational framework." They suggest that there are two determinants of language perceptions: standardization and vitality. The important point about the standard here is that it is a variety whose norms have been codified and have become associated with social dominance. Vitality refers to the number and importance of functions served, and is clearly bolstered by the status which standards possess; it can also be a feature, however, of nonstandard varieties, given sufficient numbers of speakers and community support. Furthermore, two particularly salient evaluational categories account for most of the variance: social status (which is more or less equivalent to competence) and solidarity (roughly combining integrity and attractiveness) (see Edwards, 1992).

These findings are interesting for several reasons, but the single most important factor is their stereotypical nature: people are evaluated in terms of characteristics that, in a broad-brush sort of way, reflect perceptions of the group to which they are seen to belong. The implication is obvious: **individuals** – with all their personal strengths and weaknesses – are viewed in stereotypical **group** terms. Studies have suggested that at school, in the workplace, in counselor–client relations, and so on, negative stereotypes can create problems. (It is also worth remembering here that stereotyping can operate in the opposite direction, too – those with the socially 'right' attributes may have their progress unfairly expedited.)

Difference, Deficit . . . and Solidarity

In the Classroom

More than three decades ago, John Gumperz and Eduardo Hernández-Chavez made the following observation: "Regardless of overtly expressed attitudes . . . teachers are quite likely to be influenced by what they perceive as deviant speech . . . thus potentially inhibiting the students' desire to learn" (Gumperz and Hernández-Chavez, 1972: 105). Obviously, like other members of society, teachers have language attitudes, and it is not surprising to find that, on the whole, they have downgraded nonstandard varieties and, by extension, the capabilities of their speakers. Trudgill (1975: 63) noted, for example, that teachers were not averse to telling (some of) their pupils that their speech was "wrong . . . bad . . . careless . . . sloppy . . . slovenly . . . vulgar . . . gibberish." A study conducted in Nova Scotia (Edwards and McKinnon, 1987: 335) found similar teacher attitudes: one said that poor children were unable to "articulate their thoughts"; another observed that "Blacks have a slang language all their own – they will not use proper English when the opportunity arises."

Teachers' points of view are, in one sense, merely a particular illustration of a much broader relationship between linguistic variation and social attitudes. But in another sense, their beliefs are special – for the obvious reason that negative evaluations of children's speech styles may lead to real problems, in the classroom and beyond it. It is a cruel irony that socially disadvantaged children, who clearly struggle under all sorts of very real burdens, should be weighed down still more by evaluations of the inadequacies of their language – when these evaluations are typically based upon inaccurate assumptions. The negative reactions to their speech patterns do, however, reveal how hollow a purely academic sense of 'inaccurate assumptions' can be, in a world where attitudes can transform difference into deficit.

'At Home'

Given the strength and pervasiveness of language attitudes, there are strong tendencies for speakers of nonstandard varieties to accept, themselves, the

evaluations of others (the so-called minority group reaction is an example here; see Lambert *et al.*, 1960). While the sense that one's own speech is not very 'good' may be a common phenomenon, it is nonetheless a disturbing one – as the description above of the educational arena makes clear. Thus, Halliday (1968: 165) noted that

> a speaker who is made ashamed of his own language habits suffers a basic injury as a human being; to make anyone . . . feel so ashamed is as indefensible as to make him feel ashamed of the colour of his skin.

Beyond the obvious recommendations for greater linguistic awareness and sensitivity, one might also ask just why low-status speech varieties continue to exist. If they are generally considered inferior, why wouldn't speakers try to eradicate them? Why wouldn't language or dialect shift be a more popular option? Nonstandard speakers are hardly without adequate models for language alteration, after all. Nowadays, more than ever before, the ubiquity of the broadcast media means that virtually all have at least a passive awareness of standard forms (and research has shown how easy it can be for this to be translated into a more active quantity) (Edwards, 1989). Indeed, many years ago it was predicted that a leveling of variant speech styles would be the inevitable consequence of the wired world. But it seems not to have happened.

We have to bear in mind here that the solidarity function of **any** common language is powerful; even a low-prestige variety can act as a bonding agent, reinforcing group identity (see the focused discussion in Ryan, 1979). Group identity and its (linguistic and other) markers are known quantities and, in that sense, 'safe'; on the other hand, attempts to alter speech styles may be risky. Failure may lead to marginalization – a sense of not being a full member of any social group. Even success may prove too costly. A Mexican American who abandoned Spanish was labeled a *vendido* (Carranza and Ryan, 1975), a 'sell-out,' a defector to the other side. A French Canadian in analogous circumstances was a *vendu*. Cockney grandparents once told their grandchildren that, if they tried to speak in a 'la-di-da' accent, their mates would call them 'queer' (Bragg and Ellis, 1976). Whatever the current status of these well-known if somewhat dated cases, it is clear that the principle remains powerful. Finally here, nonstandard speakers may also refrain from attempting to alter their language patterns if the exercise is considered to be pointless – as it may be seen to be for those who are members of visible minority groups, for instance. If reactions to Black English are reflections of social evaluations of Black speakers, then learning and using some more standard form may be a waste of time. The social difficulties faced by Blacks occur largely because they are black, not because they do not use standard English (consider, for example, that there are many nonstandard-speaking Whites in comfortable circumstances).

Language Attitudes and Language Acquisition

One important aspect of attitude study has been its connection with the learning of second (and subsequent) languages. Positive attitudes, it is generally thought, are likely to facilitate second language acquisition, although it is realized that variations in the context and the perceived functions of the new medium are also important. Indeed, this has very much been the received wisdom, and there is a large literature on attitude and motivation in language learning. John Macnamara, however, appeared to take an opposing view, asserting that attitudes were of little importance in language learning (Macnamara, 1973). His argument is worthy of some attention here, since it leads to some finer-grained attention to language attitudes generally (see also Edwards, 1995).

Macnamara noted that necessity may overpower attitude – and this is clearly true. The adoption of English by the Irish, for instance, was a shift not generally accompanied by favorable attitudes; indeed, most historical changes in language use owe much more to socioeconomic and political pressures than they do to attitudes *per se*. But perhaps attitudes of a sort – instrumental attitudes – do play a part in language shift. A mid-19th-century Irishman may have hated English and what it represented, while still acknowledging the necessity and long-term usefulness of the language. A pragmatic or 'instrumental' motivation need not imply the deeper, 'integrative' motives so dear to the hearts of teachers keen on introducing their charges to new languages **and** new cultures. Similarly, there may be a useful distinction here between positive and favorable attitude; to remain with the Irish example, we might say that the attitudes towards the learning of English were positive and instrumental, but not necessarily favorable or integrative. (It must be acknowledged, of course, that instrumental–integrative and positive–favorable distinctions may change or disappear as language shift develops.)

Macnamara went on to contend that pupils' language-learning attitudes at school were also relatively unimportant. Language in the classroom, he said, was typically an unreal and artificial business, with communication subordinated to narrow academic

concerns that fail to engage children's interest. It is this situation, and not pupils' attitudes, that underlies the traditionally poor results of school language programs.

Now, it is undoubtedly true that a great failing in language classrooms has been the absence of realistic usage, but I do not think that this implies that attitudes are of small importance. The argument that the classroom is artificial is essentially a condemnation of approaches and practices, and does not of itself indicate that attitudes are trivial. In fact, attitudes may be of considerable importance precisely because of 'artificiality' – i.e., where a context is **not** perceived as pertinent to real life, or is **not** seen to be necessary, attitudes may make a real difference. The importance of favorable attitudes, then, may vary inversely with real linguistic necessity (Edwards, 2001).

Details and Directions

Within social psychology, 'attitude' is a disposition to react favorably or unfavorably, a disposition with three components: feelings (the affective element), thoughts (the cognitive element), and a resultant tendency to action (the behavioral element). One knows or believes something, has some emotional reaction to it, and then does something about it. Two points, however, should be made.

There is often considerable inconsistency between attitudes and actions. In a classic study, a Chinese couple (accompanied by the experimenter) toured the United States in the early 1930s. Visiting about 250 hotels and restaurants, they were refused service only once; however, when the investigator later wrote to the places visited, he found that more than 90% said they would not serve Chinese people (LaPiere, 1934). There are, of course, many possible reasons for this finding: immediate self-interest, a desire to avoid embarrassment, a difference between reactions to an ethnic group abstraction and assessments in concrete instances, and so on. But the inconsistency remains.

The other point is that there is often confusion between attitude and belief – a confusion particularly noticeable within the language attitude literature. For example, the answer to a questionnaire item such as 'Is a knowledge of French important for your children, yes or no?' reflects a belief. To gauge attitude would require further probing into respondents' feelings about the expressed belief. It would clearly be incorrect to assume that the answer 'yes' implied a favorable attitude: the informant could grudgingly accept that French was important for career success, while loathing both the language and its culture.

More attention to these distinctions – apart from simply improving accuracy – might be salutary in terms of understanding why evaluations take the forms they do. Given that unfavorable reactions to nonstandard varieties have typically been discussed in terms of status or prestige differentials, for instance, it could be useful to confirm this, from the respondents' point of view, by asking them the bases for their evaluations. Such an exercise might be of special interest since – as has been mentioned above – nonstandard varieties elicit positive ratings along some dimensions. It is surely reasonable, after all, to gather as much information as we can from the actual participants in language studies, as well as imposing more purely theoretical interpretations upon their responses (Edwards, 1995).

Language attitude study should perhaps be prepared to expand its scope. Various methodologies have produced a sizable body of evidence bearing on social perceptions, stereotypes, and language attitudes. We can now predict with some confidence what sorts of reactions are likely when people hear varieties of (for instance) Black English, Newfoundland English, Cockney, Received Pronunciation, Boston English, and many others; we can also make predictions about those 'ethnic' varieties produced by nonnative speakers of English that show the influence of the first language. We understand, at a general level, how these reactions come about, via the linguistic 'triggering' already noted, and how they reflect a set of attitudes (or beliefs) – often of stereotypical nature – that listeners have of speakers. Investigators have not, however, gone very much beyond fairly gross explanations; that is, they have typically not related speech evaluations to particular speech attributes. Thus, although hundreds of experiments have revealed negative reactions towards Black English, we have virtually no information relating specific linguistic attributes of that variety to such reactions – attributes that might include pronunciation patterns, particular grammatical constructions, or the use of dialect-specific lexical items (or, of course, any combination of these) (Edwards, 1999).

Some moves have already been made in this direction. For example, in 1982, Howard Giles and Ellen Bouchard Ryan issued a call for "more detailed linguistic and acoustic descriptions of the stimulus voices as well as examining the relative evaluative salience of these particulars for different types of listeners" (Giles and Ryan, 1982: 210). A related, although not so pointed, observation was made by Edwards, in calling for fuller probing of the reasons behind the evaluative decisions made by informants (Edwards, 1982; see also above). In general, though, social psychologists have done little in the way of

isolating linguistic and acoustic variables and relating them to evaluative judgments. Such work has simply not been their métier.

It is in linguistic research that we find descriptions of features that characterize and differentiate language varieties. In recent years, a considerable amount of work has been done here; for example, linguists (including Laver and Trudgill, 1979 – in a book on social markers that remains a valuable reference) have pointed out such phenomena as:

1. the nasality habitually associated with some (RP, for example) but not all varieties of English;
2. the wide dialectal variations in consonant pronunciation: thus, RP speakers pronounce *lock* and *loch* more or less identically, with a final /k/, but (some) Scottish pronunciations involve final /x/. Another example: in British English, pronouncing the postvocalic /r/ in words like *cart* and *mar* is inversely related to social-class status, whereas in some varieties of American English (e.g., New York English) a positive correlation exists between /r/ pronunciation and status;
3. grammatical variation: copula deletion in American Black English, for example (as we have seen, above);
4. lexical differences: some English speakers *brew* their tea, some *mash* it, some let it *steep,* some let it *set,* and so on.

If, however, linguists have been the ones to describe such variation, they have either been relatively uninterested in its relation to variations in social evaluations, or have simply assumed that the more obvious and salient linguistic markers are the triggers for differentiated ratings. Like social psychologists, linguists too have generally stuck to their lasts – although a recent and most welcome collection (Milroy and Preston, 1999) deals explicitly with linkages between linguistic features and attitudes.

Conclusions

There are many important aspects of language attitudes and their ramifications that it has not been possible to attend to in this brief overview. One has to do with the relationship between language and gender. The 'covert prestige' briefly noted above, for instance, is essentially a male phenomenon – the positive connotations of nonstandard speech are not typically found among women; indeed, where middle-class male informants may report more nonstandard usage than is actually the case, female respondents overreport standard dialect usage. A large literature suggests, generally, that women are often more linguistically conservative than are men, that they produce 'politer' and more 'correct' speech, that there are significant gender differences in lexicon and function, and so on. All of these features are influenced by, and reflected in, language attitudes (see Coates, 1993 for a good overview).

Another important area with a burgeoning literature has to do with the accommodations made by speakers in different contexts and with different interlocutors. Linguistic accommodation can take many forms but, whether it operates at or below the level of conscious awareness, its fundamental feature is the modification of speech patterns to converge with, or diverge from, those of others. Accommodation can reflect individual concerns – wanting to sound more like the boss, or intentionally departing from the usage of someone you dislike – or group ones: you may wish to emphasize your 'in-group' membership, or to solidify an ethnic or class boundary. Once again, attitudes clearly underpin practices (for an overview here, see Giles and Coupland, 1991; Robinson and Giles, 2001).

What I have tried to deal with in this article is some of the central themes in the area, some of the scaffoldings on which other aspects rest. They include the following:

1. some basic definitional and terminological matters: it is always important to be clear about terms and meanings, and especially so when – in areas such as this – the terms have broad popular connotations as well as more narrowly academic ones;
2. the bases upon which linguistic judgments are made: most people still believe that intrinsic qualities of language varieties can reasonably underpin rankings of 'better' and 'worse.' The persistence of this belief is disappointing, to say the least, especially in educational circles;
3. the manner in which linguistic difference is typically translated into linguistic deficit: it is evident that, linguistic enlightenment notwithstanding, society continues to deal rather bluntly with variation – and, if most consider a difference to be a deficit, then it becomes one;
4. the ways in which linguistically stigmatized individuals and groups come to accept the judgments made of them by 'outsiders': this is an extension of the previous point, insofar as it is a potent reflection of the power of social pressure;
5. the reasons behind the persistence of low-prestige varieties: given the potency of social pressure and prejudice, it is interesting that group solidarity, on the one hand, and fear of personal marginalization, on the other, are sufficient to maintain stigmatized varieties;

6. the evaluative dimensions and categories that seem to be employed when ranking different language varieties: the regularity with which a few basic categories (such as 'competence,' 'integrity,' and 'reliability') appear across a very wide range of studies is one of the most robust features in language study;
7. language attitudes in contexts of learning and shift: questions of how, and to what extent, attitudes and motivations are important in situations of change – or potential change – throw light upon both the attitudes and the contexts;
8. the need for greater interdisciplinarity in what are sometimes called, hopefully, the 'language sciences': a common criticism of modern social psychology involves its allegedly disembodied character – further and deeper attention to language, and language attitudes, would at once redress academic oversight and inevitably abet interdisciplinary linkages that would, in turn, reduce decontextualization.

See also: Communities of Practice; English, African American Vernacular; Ethnicity; Ethnolinguistic Vitality; Interactional Sociolinguistics; Minorities and Language; Multiculturalism and Language; Social Class and Status.

Bibliography

Bragg M & Ellis S (1976). *Word of mouth*. London: BBC.

Carranza M & Ryan E (1975). 'Evaluative reactions of bilingual Anglo and Mexican American adolescents toward speakers of English and Spanish.' *International Journal of the Sociology of Language 6,* 83–104.

Chapman R W (1932). 'Oxford English.' *(Tracts of the) Society for Pure English 37,* 540–562.

Coates J (1993). *Women, men and language* (2nd edn.). London: Longman.

Edwards J (1982). 'Language attitudes and their implications among English speakers.' In Ryan & Giles (eds.) 20–33.

Edwards J (1989). *Language and disadvantage* (2nd edn.). London: Cole & Whurr.

Edwards J (1992). 'Language in group and individual identity.' In Breakwell G (ed.) *Social psychology of identity and the self concept*. London: Academic Press. 129–146.

Edwards J (1995). *Multilingualism*. London: Penguin.

Edwards J (1999). 'Refining our understanding of language attitudes.' *Journal of Language and Social Psychology 18,* 101–110.

Edwards J (2001). 'Languages and language learning in the face of world English.' *Profession* 109–120.

Edwards J & McKinnon M (1987). 'The continuing appeal of difference as deficit: a Canadian study in a rural context.' *Canadian Journal of Education 12,* 330–349.

Giles H & Coupland N (1991). *Language: contexts and consequences*. Milton Keynes: Open University Press.

Giles H & Ryan E B (1982). 'Prolegomena for developing a social psychological theory of language attitudes.' In Ryan & Giles (eds.) 208–223.

Giles H, Bourhis R, Trudgill P & Lewis A (1974). 'The imposed norm hypothesis: a validation.' *Quarterly Journal of Speech 60,* 405–410.

Giles H, Bourhis R & Davies A (1979). 'Prestige speech styles: the imposed norm and inherent value hypotheses.' In McCormack W & Wurm S (eds.) *Language and society: anthropological issues.* The Hague: Mouton. 589–596.

Gumperz J & Hernández-Chavez E (1972). 'Bilingualism, bidialectalism and classroom interaction.' In Cazden C, John V & Hymes D (eds.) *Functions of language in the classroom.* New York: Teachers College Press. 98–111.

Halliday M (1968). 'The users and uses of language.' In Fishman J (ed.) *Readings in the sociology of language.* The Hague: Mouton. 160–172.

Labov W (1976). *Language in the inner city.* Philadelphia: University of Pennsylvania Press [includes his famous 1969 article, 'The logic of nonstandard English'].

Lambert W, Hodgson R, Gardner R & Fillenbaum S (1960). 'Evaluational reactions to spoken languages.' *Journal of Abnormal and Social Psychology 60,* 44–51.

LaPiere R (1934). 'Attitudes versus actions.' *Social Forces 13,* 230–237.

Laver J & Trudgill P (1979). 'Phonetic and linguistic markers in speech.' In Scherer K & Giles H (eds.) *Social markers in speech.* Cambridge: Cambridge University Press. 1–32.

Macnamara J (1973). 'Attitudes and learning a second language.' In Shuy R & Fasold R (eds.) *Language attitudes: current trends and prospects.* Washington: Georgetown University Press. 36–40.

Milroy L & Preston D (eds.) (1999). *Attitudes, perception and linguistic features:* special issue of *Journal of Language and Social Psychology 18.*

Robinson P & Giles H (eds.) (2001). *The new handbook of language and social psychology.* London: John Wiley.

Ryan E (1979). 'Why do low-prestige varieties persist?' In Giles H & St. Clair R (eds.) *Language and social psychology.* Oxford: Blackwell. 145–157.

Ryan E B & Giles H (eds.) (1982). *Attitudes towards language variation.* London: Edward Arnold.

Ryan E B, Giles H & Sebastian R (1982). 'An integrative perspective for the study of attitudes toward language variation.' In Ryan E & Giles H (eds.) 1–19.

Ryan E B, Hewstone M & Giles H (1984). 'Language and intergroup attitudes.' In Eiser J (ed.) *Attitudinal judgment.* New York: Springer.

Trudgill P (1975). *Accent, dialect and the school.* London: Edward Arnold.

Trudgill P (1983). *On dialect.* Oxford: Blackwell.

Wilkinson A (1965). 'Spoken English.' *Educational Review 17*, supplement.

Wyld H (1934). 'The best English: a claim for the superiority of received standard English.' *(Tracts of the) Society for Pure English 39,* 603–621.

Language Change and Cultural Change

J H Hill, University of Arizona, Tucson, AZ, USA

Human languages are inherently variable and dynamic, and their shifting shapes are constrained by diverse factors. Culture, defined as those human projects which endow the raw materials of our world – including the stuff of language – with meaning, must be among these. But to explore relationships between language change and cultural change requires appreciation of the immense complexity of cultural mediation, and linguistic analysis at a high level of detail and delicacy.

Several basic observations are required prior to developing this theme. First, 'languages' do not map isomorphically onto 'cultures.' From this obvious fact it is immediately clear that the relationship between language change and cultural change will not be straightforward. Mutually intelligible varieties of large regional and world languages are found among people involved in very diverse ways of life. Mutually intelligible varieties of Spanish are spoken by peasants in Mexico and philosophers in Madrid. Conversely, people who share very similar ways of life may speak several different languages. The Pueblo Indians of the southwest United States constitute a highly coherent and relatively homogeneous culture area, with a long history of contact and exchange among their communities, but speak at least five mutually unintelligible languages belonging to three different language families, plus one isolate. Local language ideologies work to minimize convergence among these languages (Kroskrity, 1993).

Second, resonances between language change and cultural change tend to be quite local. Linguists correctly greet global generalizations with profound suspicion. For instance, there is no universal correlation between the elaboration of sexism and gender-based discrimination in a society and the grammaticalization of gender in its language or languages. In an exceptionally careful study of this question, Chafe (2002) explored the relationship between gender roles and the grammar of gender in Northern Iroquoian. In several languages of the group, a set of masculine pronominals distinguishes singular and plural forms and subject and object case. A feminine gender pronominal has developed from an original 'nonspecific' pronoun, and in the nonsingular is not distinguished from neuter gender. Chafe argued that contrastive elaboration of masculine gender versus feminine gender pronominals, which occurred about a thousand years ago, is in resonance with the prominence given to males in many Northern Iroquoian cultural scenarios (although Northern Iroquoian societies are matrilineal). However, he noted that such male prominence is extremely common in human communities, yet most languages do not show the same kind of differential elaboration of gender marking. For this reason, he preferred to speak of a cultural 'motivation' for the pattern, rather than of a cultural 'cause.'

In the Northern Iroquoian example, the structure of the pronominal system is unlikely to be prior to, and motivating of, the prominence of male roles. Instead, male prominence motivated the elaboration of the masculine pronouns. The diverse developments of noun classification through plural marking in the Uto-Aztecan languages (Hill and Hill, 1997) provide another case of the cultural motivation of linguistic patterning, in this case a loss of elaboration. In Proto-Uto-Aztecan, reduplicated plurals marked a class of human nouns, as against other nouns pluralized with suffixes. In two descendant language groups, the O'odham or Upper Piman languages and Hopi, the marked class has been extended beyond human nouns to classify for 'protrusion' and 'intrusion.' In Tohono O'odham (a Uto-Aztecan language of Arizona), spoken in a community that observed patrilineal descent, lineage exogamy, and patrilocal residence, the marked plural class includes kinship terms for women and kin linked to ego by women, but not kinship terms for men. This resonates with local understandings, expressed in legends, that inmarrying women stand as marked outsiders to the patrilineal kin group. In contrast, in Hopi, with matrilineal descent and matrilocal residence, the only marked plural kin terms are the words for 'bride, inmarrying woman' (an outsider from the point of view of the husband's matrilineage) and 'husband, inmarrying man' (an outsider from the point of view of the wife's matrilineage). That is, the metaphor of 'outsiderness' functions in the noun classification system in the kin terms of both languages, but is greatly restricted in Hopi, consistent with the understanding there that descent through females constitutes the core social unit, the matrilineage. It seems most unlikely that patterning in the pluralization of kin terms would shape social organization, but quite reasonable that understandings of social organization would motivate the use of marked versus unmarked plural forms.

The loss in English of the distinction, otherwise nearly universal in European languages, between a polite/distant and an intimate/solidary second-person pronoun provides another example of such culturally motivated language change. Most scholars believe that the English T-pronouns, the intimate/solidary

second person forms *thou*, *thee*, *thy*, *thine* were lost in ordinary usage in the late 17th century in reaction to the linguistic practices of a dissident religious sect, the Society of Friends or Quakers. Friends, consistent with their belief in the equality of all human beings before God, refused to use the V-pronouns that signaled social distance and deference. This practice was scandalous in a society where a regard for hierarchy was hegemonic and the use of a T-pronoun could constitute insult or even illegal *lèse majesté* that might lead to imprisonment. Rather than risk being taken for a Quaker, other speakers displayed their commitment to hierarchy and propriety by generalizing the V-pronouns even to contexts that had hitherto called for the use of the T-form without any presupposition of insult or disrespect. Regardless of whether the role of Quaker innovation has been properly understood, it seems likely that cultural ambivalences and ambiguities about hierarchy in the emergence of modernity shaped the gradual loss of the T/V distinction in English, rather than that the loss of the T/V distinction motivated a reorganization of British social order.

Benjamin Whorf, often misunderstood as a linguistic determinist, outlined in his most significant paper (Whorf, 1956) a dialectical interaction between language change and cultural change that attempted to account for examples like those above. Silverstein (1976, 1979, 1996) has proposed a contemporary semiotic refinement of Whorf's ideas. He argues that general semiotic principles underlie the development of such resonances between language change and culture change. These principles produce a 'dialectic of indexicality' which drives analogical change. First, speakers project referential functions on to indexical functions and 'objectify' these. In the case of the English (and other European language) pronouns reviewed above, Silverstein suggests the following chain of analogies. The third-person pronouns exhibit apparent referential distinctions of number and gender, which are taken by speakers to refer properly to things in the world – although these 'things' achieve their 'naturalized' status, of course, precisely through linguistic practice and specifically through indexicality. But speakers come to see this indexical projection as 'reference.' By analogy, the social hierarchy distinctions of the T/V contrast in the second-person pronouns, which are technically indexical, can be taken by speakers to be referential as well. Thus to English speakers the social world came to be seen as consisting of individuals possessing an inherent property of meriting deference, parallel to inherent properties of gender, or of singularity versus plurality. Such analogical reasoning permits speakers to manage social anxieties (Inoue, 2004), by trying to bring usage rigidly into line with an understanding of the world that has precisely been projected from language in the first place. Such 'iconizing' (Gal and Irvine, 1995) projects do not always succeed in taking a language change to completion, but it is clear that the constant interplay of indexical orders (Silverstein, 1996), motivated by specific cultural anxieties and ambivalences that must be ever present in human communities, constitutes a powerful force for language change.

Among the kinds of change that can be driven by the 'dialectic of indexicality' is the formation and spread of styles and registers. An excellent example is the shift in American English in the stylistic features thought to be appropriate to the talk and text of ministers, political orators, and the like in their public roles. Cmiel (1990) has shown that an explicit ideology of egalitarianism emerging through the end of the 18th and the first half of the 19th centuries led to the development of the so-called middling style for elevated language. In the middling style ordinary vernacular words and short sentences came to be favored. This process of reshaping public language as an icon of the American egalitarian imaginary continues to develop. The grammar and lexicon of documents such as the Declaration of Independence and the U.S. Constitution strike many 21st-century Americans as forbiddingly elevated and elite. Even liturgical language has been simplified: new translations of the Christian Bible are in strikingly everyday language, and ministers of the gospel experiment with vernacular forms (as in the famous case of Cornell West, a minister and professor at Harvard University, who recorded a compact disk of raps, a genre that originated in African-American street culture). Highly placed American politicians can use quite vulgar language, even in elevated public contexts, without facing censure – indeed, such language may bring approbation. It marks them not only as ordinary people with a commitment to egalitarianism, but as plain-speaking 'straight shooters.' It is perhaps not surprising that the tendency toward 'plain speech' in public forums has reached new heights (or depths) at precisely a period when increasing political-economic inequality has made many Americans anxious about the status of egalitarianism, a continuing ultimate value in U.S. society.

Language Change and Sociocultural Evolution

The work briefly reviewed above examines very fine-grained linguistic changes motivated by ideological processes in local speech communities. However, linkages between more general tendencies of

language change and cultural change have often been proposed. Widely accepted through the early 20th century was the theory that human language exhibited a hierarchy of evolutionary grades, from 'primitive' to 'civilized,' that correlated with the grades of the theory of cultural evolution. While linguists today almost universally reject such ideas, evolutionary thinking continues to have a place in theories of the relationship between language change and culture change, and deserves brief attention.

Certain limited features of human languages may be loosely correlated with subsistence type and social complexity. These include plant and animal terminologies. For instance, Brown (1984) argued that the number of life-form terms for plants (*tree*, *bush*, *herb*, *grass*, *vine*, etc.) is correlated with increasing social complexity. Brown also suggested that cultivators have many more folk generic terms for plants than do hunter-gatherers, and also use more binomial species labels. While other studies have supported these results (e.g., Balée, 1994), they should be interpreted with caution. Balée has pointed out that the difference in the number of generics for plants between cultivators who speak Ka'apor (Urubú-Kaapor) and hunter-gatherers who speak Guajá, another Tupi-Guaraní language closely related to Ka'apor, is due partly to the fact that the Ka'apor simply have more uses for plants; they use implements made from plants to process other plants, and their sedentism permits an elaborate material culture. Furthermore, Hunn (2001) has pointed out that in Oaxaca, vernacular speakers of Spanish, an expanding world language, have far simpler repertoires of biosystematic terminology than do speakers of small local Zapotecan languages, in spite of the fact that Spanish has been spoken in the area for 500 years. Berlin and Kay (1969) held that elaboration of color terminologies was correlated with level of sociocultural evolution. But MacLaury (1991) showed in a study of Mayan-language speakers in Chiapas that the elaboration of color terminology could exhibit vast differences among speakers of a single language, and seemed to depend more on the specific life circumstances of individuals, especially on how much people were forced to engage with novelty, than on any broad correlation between language and sociocultural evolution. Thus terminological elaboration in the biosystematic and color lexicons appears to be largely correlated with the elaboration of attention and use, a matter that is not surprising in any way and that can be accounted for within Silverstein's model outlined above, without appeal to evolution theory.

The concept of natural selection, taken from Darwinian theory, has been invoked by modern linguists to account for the relationship between language change and cultural change. Hymes (1968) suggested that the current reduction of world linguistic diversity, caused by the expansion of 'world' languages at the expense of 'full' languages, results from a selective process in which world languages are better adapted to certain emerging communicative functions, and thus replace full languages. Croft (2000) and Mufwene (2001) have suggested that linguistic elements within inherently variable assemblages (which Mufwene has characterized as 'Lamarckian species') – phonetic realizations, construction types, semantic features, and the like – can be seen, like genetic alleles, as targets of natural selection. Cultural context, among other variables, yields selective pressures, and the relative frequency of particular realizations of linguistic elements will reflect these. The development of Haitian Creole (Haitian Creole French) from French has been treated within this model. During the early 'settler' period, when slaves lived intimately with their French masters on small landholdings, they acquired a version of vernacular French that, while it was a second language, was not greatly dissimilar from the French spoken by the native speakers from whom they learned it. The shift to plantation agriculture, however, swamped out these early second-language varieties of French. Increasingly large populations of newly imported African slaves had less and less access to French models. Themselves learning the language from other second-language speakers, they favored the least marked French constructions and constructions derived from African sources over the varieties used by the earlier African settlers of Haiti. This process yielded the basilectal creole varieties attested in Haiti today. Thus, the context for selection that favored basilectal structures was a cultural system: plantation agriculture, with its huge slave work forces supervised by relatively few – and interactionally remote or absent – French-speaking overseers.

Advocates of the selectionist view have neglected the complex mediations between cultural and linguistic systems that must underlie the emergence of a system such as Haitian Creole, and that are treated in the semiotic theory of Silverstein and others cited above. Constraints on the acquisition of second-language features do not consist solely of the opportunity for learners to observe native-speaker models, and such opportunity is not determined in any simple way by mode of subsistence or labor organization. Instead, these opportunities and what learners make of them are mediated by ideological systems. For instance, in the Haitian Creole case European French speakers almost certainly found that the shift toward

basilectal varieties was entirely consistent with their understanding of Africans as unintelligent and unable to master refined European speech. With such an understanding – a convenient one when the labor regime involved starving and torturing these Africans – the 'disorderly' grammar and 'impoverished' lexicon of the slaves would be merely natural, would invite no intervention, and might even be encouraged. Deviations from this 'iconization' would, in Gal and Irvine's (1995) terms, be 'erased,' and pass unnoticed or be dismissed as aberrant or suppressed as dangerous. At the same time, among Africans, basilectal varieties could become icons of their own world, symbols of resistance to and differentiation from their French-speaking oppressors. An approach restricted to the terms of the theory of natural selection does not invite investigation of the semiotic dynamic by which these ideological systems – which are well documented – resonated with the emerging basilectal structures.

The development of modern Mexicano in the state of Tlaxcala, Mexico shows clearly that simple demographic facts cannot account for language change in a contact system. Mexicano is the descendant of Tlaxcalan Nahuatl (Náhuatl Central), one of the Nahua (Aztecan) languages of the Uto-Aztecan family. Mexicano has incorporated very large amounts of Spanish loan vocabulary, across nearly every subcomponent of the lexicon from nouns for European cultural elements such as horses and chickens to discourse particles, prepositions, and derivational morphemes. The Mexicano case is the reverse of the Haitian Creole case: over a few hundred years, immigration from Europe reduced Nahua-speaking communities to islands in a Spanish-speaking sea. However, the rate of incorporation of Spanish loan elements into Mexicano cannot be correlated directly with the ratio of Nahuatl speakers to Spanish speakers. Instead, it is closely linked to genre. Spanish loan words can convey an elevated or serious tone, and so large numbers of Spanish constructions appear even in the 16th century, at a time when Nahuatl speakers were emphatically the majority in central Mexico and very few people were bilingual, in such genres as the minutes of the proceedings of town councils and notarial records such as wills. Bilinguals today can adjust the frequency of Spanish versus Nahuatl roots even in spontaneous speech. They express thereby subtle variations in stance and attitude that are associated with diverse social positions, age cohorts, and genders within a larger ideological system that contrasts nativistic Nahuatl purism and the expression of community solidarity with an ideology that, by iconization, links Spanish to the power of its first-language speakers.

Language Spread and Language Shift

Cultural mediation appears to play a role in language spread and language shift. The archaeologists Colin Renfrew and Peter Bellwood (2002) hypothesized that large multibranched language families such as Indo-European or Niger-Congo-Kordofanian almost all descend from languages spoken by primary cultivators at the time of the initial domestication of animals and cultivated plants. In contrast, small language families such as Wakashan or North Caucasian are the residue of a more linguistically diverse pre-Neolithic epoch. Renfrew and Bellwood argued that the principal mechanism of language spread throughout most of human history was not the language shift that is so obvious in the history of the last 500 years, but the expansion and spread of populations of speakers of the languages ancestral to the large families. Since demographic factors favor the increase of populations of cultivators over populations of hunter-gatherers, Renfrew and Bellwood predicted that the languages of the latter will be found only in relic zones where the Neolithic expansions of cultivator populations did not reach.

One difficulty with this very general model is that there are attested cases of ancient language shift. For instance, two of the major language families of the Americas, Athabaskan and Algonquian, are spoken largely by hunter-gatherer populations which obviously expanded without shifting to cultivation. Some of the descendant languages (such as California Athabaskan, Denaina in southwest Alaska, and the Algonquian descendant language Cree-Montagnais [Montagnais] in the Canadian subarctic) exhibit incongruent elements that must represent the influence of speakers of a substratum language who adopted the expanding Athabaskan or Algonquian variety (Kari, 1989; Golla, 2000). The Australian linguists Nicholas Evans and Patrick McConvell (Evans and McConvell, 1998; McConvell, 2001) have proposed that hunter-gatherers who survive climatic instability in continental interiors must be able to accomplish not only movements into uninhabited territories, but retreats into populated regions. The latter type of move requires the cultural capacity – Evans and McConvell call it 'charisma' – to incorporate populations thus encountered into the social, ritual, and linguistic systems of the retreating group. Such movements are attested by apparent substratum linguistic elements not only in the Americas, but in the Australian materials which Evans and McConvell originally investigated. Evans and McConvell's work suggests that a set of ideas about language must be as much a part of successful hunter-gatherer adaptations as is nature knowledge and technological capacity.

Nichols (1992) suggested that the likelihood of language shift is conditioned by human ecology. She proposed a distinction between linguistic 'accretion zones' (sometimes called 'residual zones') and linguistic 'spread zones.' The former tend to form in micro-environmentally complex coastal and montane zones, and are characterized by very large numbers of languages, large numbers of small language families, and high levels of multilingualism. Indigenous New Guinea, aboriginal California, and the Caucasus mountains are examples of accretion zones. Nichols argued that the languages of accretion zones are relatively stable and thus are likely to harbor archaic features. 'Spread zones' tend to form in continental interiors and are characterized by relatively small numbers of languages and language families, and the formation of lingua francas. Languages in spread zones are more likely to be innovative and fail to preserve archaic features. The Eurasian steppes, the Congo basin, and the North American subarctic are examples of spread zones. Language shift, to spreading languages, is much more likely in spread zones than in accretion zones. At the root of this distinction is probably the differential cultural adaptations of humans to the two different types of environment. In the intricate mesh of relatively stable micro-environments in accretion zones such as aboriginal California or the Caucasus, cultural systems that favor local differentiation and specialization, mediated by various forms of cultural exchange, tend to be associated with linguistic ideologies that favor multilingualism over language shift. Local ways of speaking become markers of privileged access to local resources, but speakers become multilingual to manage exchange relationships. In contrast, in spread zones a combination of relative environmental instability and large-scale homogeneity of human adaptive strategies may favor distribution of risk over large geographical regions. This is best accomplished on the linguistic side through the relative homogeneity of a spreading language. A dramatic example of this type is the spread of Shoshoni across the Great Basin of North America. Shoshoni language ideology favors inattention to dialect variation, this inattention being accompanied by an unusually high level of mobility of bands and family groups, and Shoshoni exhibits no clear linguistic isoglosses even though it is distributed over an extremely large area (Miller, 1970, 1971). It seems likely that this sociolinguistic picture is an old one in the community. However, language shift in accretion zones **is** attested, as with the language shift in California and coastal Alaska to California Athabaskan and Denaina respectively, with the superstratum Athabaskan communities being involved in an Evans and McConvell-type move into populated territory.

Hill (2000) proposed that different tendencies in the organization of language diversity between accretion zones and spread zones probably result from complex interactions between ecological adaptation, the social contexts of language acquisition, and language ideologies. When linguistic ideologies construct features of language as licensing access to resources, two directions in sociolinguistic stance by speakers are possible. The adoption of a 'localist' stance is exemplified by speakers of the Kokolo:di dialect of Tohono O'odham, who live in a relatively well-watered and productive region of the Sonoran desert. Their dialect is highly homogeneous and focused, and speakers are very resistant to, and self-conscious about, innovation. Speakers of the Gigimai dialect, on the other hand, who live in a region with sporadic and very low rainfall, exemplify a 'distributed' sociolinguistic stance. These speakers are relatively promiscuous adopters of all available linguistic resources. Hill suggested that localists understand their situation in relation to their language and their environment as entailing a statement like 'I have a rightful and primary claim on valuable and dependable local resources that are necessary to my well-being, and the way I speak verifies my claim.' In contrast, distributed speakers appear to believe that 'I have no rightful and primary claim on valuable and dependable local resources adequate to sustain my well-being. However, I might be able to add to my limited primary claims secondary claims on a sufficient range of a distributed inventory of resources to sustain my well-being. The way I speak verifies these secondary claims.' Even in this case of the desert-dwelling Tohono O'odham, where historically water scarcity approached and sometimes passed limits on human survivability, the contrasting sociolinguistic stances are not strictly determined by environment, but by local cultural understandings of it and by a linguistic ideology that takes certain sociolinguistic stereotypes to license insider versus outsider status, and hence rights to resources. The Tohono O'odham do not live in the same desert that a biologist or a geologist might see, but construct 'landscapes,' culturalized environments in which resources become meaningful. Furthermore, unlike the desert-dwelling Western Shoshoni described by Miller (1970, 1971), the Tohono O'odham choose to notice dialect differences, construct these as sociolinguistic stereotypes, and use them to license rights. There are, however, ecological constraints that can rule out a localist stance. Localism, with its close adherence to relatively homogeneous ways of speaking, probably requires a period of focusing within childhood peer groups,

and the local environment must provide enough in the way of relatively stable subsistence resources to permit the formation of such groups. Distributed-stance speakers of Tohono O'Odham report that water and food scarcity required their families to make frequent moves, precluding the formation of stable peer groups. Among these speakers, even siblings may speak quite differently from one another. Miller (1970, 1971) reported absence of child peer groups for the Western Shoshoni as well. Childhood peer groups, or lack thereof, are not the only dimension of social organization that may shape language spread and shift. James and Leslie Milroy (1985) showed for urban contexts that sociolinguistic 'focus,' the relative homogeneity and stability of dialects, is related to the structure of social networks, which is in turn related to local impacts of political-economic systems. Relatively dense and closed networks, where people have many different kinds of relationship with others that remain relatively stable, favor strong focusing. In contrast, open networks, in which people often form new relationships, some of these involving so-called 'weak ties' – social ties that mediate only very limited kinds of interaction, such as commercial transactions – seem to favor the rapid spread of innovation and relative heterogeneity and instability in local ways of speaking.

Esoterogeny and Exoterogeny

Thurston (1987, 1989, 1992), based on work in Melanesia, suggested that types of structural drift in language change may be associated with the localist and distributed stances outlined above. Very focused local varieties may undergo what Thurston called 'esoterogeny,' acquiring and conserving highly marked and unusual typological properties that are difficult to learn. In contrast, spreading languages are likely to be 'exoteric,' characterized at many loci of typological variation by relatively unmarked structures. An illustration of contrasting esoterogenic and exoterogenic processes of change can be drawn from the Uto-Aztecan languages. The ancestral Uto-Aztecan protolanguage seems to have had at least a distinction between an unmarked nominative and a marked accusative case on nouns, and a set of argument-encoding clitics that distinguished nominative from accusative. The descendant Nahua languages, which have spread throughout central Mexico and into Central America as far south as El Salvador and Honduras and had millions of speakers (Náhuatl in Mexico still has at least one million speakers), retain only the latter feature and lack case marking on nouns. In contrast, Cupeño, spoken within the California accretion zone and perhaps the smallest Uto-Aztecan language, marks the accusative case on nouns and their modifiers, distinguishes nominative from accusative case in a set of pronominal prefixes, has added the additional elaboration of a set of person-number clitics that distinguish an ergative agent from an absolutive subject, and distinguishes the absolutive objects of imperative verbs from accusative-marked objects of verbs in nonimperative moods.

One of the most famous examples of esoterogeny is the Amazonian language isolate Pirahã (Múra-Pirahã), spoken by about 150 people. Everett (2004b), the major contemporary scholar of the language, has commented that "the combination of a small phoneme inventory, which provides minimal contrast, complex morphology, tone and a few other features unique to Pirahã make this language very unusual and extremely difficult to learn." The unusually small inventory of segments includes some otherwise unattested allophones, such as a bilabial trill allophone of /b/ and a double-flap allophone of /g/, in which the tongue tip first flaps on the alveolar ridge and then on the lower lip as the tongue actually comes out of the mouth. The stress rules of the language involve an unprecedentedly complex interaction between rightward alignment and four distinct syllable weights, which are shaped partly by the voicing of the onset consonant rather than the weight of the nucleus (there are no coda consonants). Such 'esoteric' features are almost certainly propped up by local ideologies that have made them meaningful as identity markers. This is the case with Pirahã. In contrast with their close linguistic relatives, former speakers of Mura who now speak only Portuguese, and in spite of ample contact with Portuguese speakers over 200 years, the Pirahã are not bilingual, believing that speaking only their own language is essential to their identity as people who are 'straight' instead of 'crooked' (Everett, 2004a).

Culture and the Speed of Language Change

Historical linguists are very interested in the speed of language change, since methods of linguistic prehistory can retrieve deeper pasts when language change is slow. Thus, it would be useful to determine whether cultural factors can slow or accelerate language change. There is not time to treat this point in detail; a very useful discussion of factors that slow and accelerate cultural change itself can be found in Urban (2001). Therefore, I treat only two salient questions, the role of subsistence type and the role of media. In the case of subsistence type, it is often suggested that language change may be faster among

hunter-gatherers than in more complex societies. However, the above discussion has insisted that susceptibility to innovation is mediated by much more complex factors. The distinction between accretion zones, where speakers tend to adopt localist stances, be relatively resistant to innovation, and display tendencies toward esoterogeny when they do innovate, versus spread zones, where speakers tend to adopt distributed stances and be relatively open to innovation which is likely to be exoteric, crosscuts the distributions of hunter-gatherers and cultivators. In accretion zones are found hunter-gatherers such as the Karuk of California, but also communities of small-scale horticulturalists such as those of highland New Guinea, and even literate state-level societies as in the Caucasus mountains. Spread-zone linguistic communities also include societies at every level of sociocultural evolution. Thus, no generalization about the speed of language change that refers simply to subsistence type or level of sociocultural complexity is likely to be useful.

Another argument frequently made about the interaction between sociocultural factors and rate of language change is the idea that the adoption of certain media will slow down or speed up the rate of innovation. In the case of writing, often suggested as a factor favoring conservatism, this is almost certainly wrong. This can be shown by the well-known contrast between two Germanic languages, Icelandic and English, both spoken in highly literate societies. Icelandic is a notorious case of conservatism, with today's speakers able to understand 800-year-old texts with minimal assistance. In contrast, English of the same period is quite inaccessible to contemporary speakers, and must be studied as a foreign language, even though literacy has been quite widespread in England for nearly 500 years (Cressy, 1980). Recently developed media such as radio, television, and the Internet are often suggested as likely to favor accelerated innovation and the spread of languages spoken by early adopters of such media. The lesson of literacy, for which we have thousands of years of records in hundreds of languages, is that the interaction between media and language change is likely to be complexly mediated by ideology and by the details of local social organization and political economy.

Conclusion

The cases and controversies about the relationship between language change and cultural change reviewed above lead historical linguists to conclude that the two processes seem often to be connected. However, these connections seem to be irrevocably local and historical rather than general and evolutionary. Tendencies – such as the loose association between accretion zones, localist sociolinguistic stances, and esoterogeny versus spread zones, distributed sociolinguistic stances, and exoterogeny – can indeed be identified and are useful in pointing toward possibilities. But exceptions are common, and evolutionary accounts of local situations are likely to be unsatisfactory. Instead, the understanding of these relationships is a historical-linguistic project. Such a project requires detailed historicist untangling of the interaction between local cultural construal of resources, rights, and persons and the ways that local details of drift in language change are linked to these by language ideologies.

See also: Cultural Evolution of Language; Environment and Language; Evolutionary Theories of Language: Current Theories; Evolutionary Theories of Language: Previous Theories; Indexicality: Theory; Language Change, Intentional; Language Ideology; Language Maintenance and Shift; Language Spread; Media and Language: Overview; Metasemiosis and Metapragmatics; Nahuatl; Origin and Evolution of Language; Primitive Languages.

Bibliography

Balée W (1994). *Footprints of the forest: Ka'apor ethnobotany – the historical ecology of plant utilization by an Amazonian people.* New York: Columbia University Press.

Berlin B & Kay P (1969). *Basic color terms.* Berkeley: University of California Press.

Brown C (1984). *Language and living things.* New Brunswick, NJ: Rutgers University Press.

Chafe W (2002). 'Masculine and feminine in the Northern Iroquoian languages.' In Enfield N J (ed.) *Ethnosyntax.* Oxford: Oxford University Press.

Cmiel K (1990). *Democratic eloquence: the fight over popular speech in nineteenth century America.* New York: W. Morrow.

Cressy D (1980). *Literacy & the social order: reading & writing in Tudor & Stuart England.* Cambridge: Cambridge University Press.

Croft W (2000). *Explaining language change: an evolutionary approach.* London: Longman.

Evans N & McConvell P (1998). 'The enigma of Pama-Nyungan expansion in Australia.' In Blench R & Spriggs M (eds.) *Archaeology and language 2: archaeological data and linguistic hypotheses.* London: Routledge.

Everett D (2004a). 'Cultural constraints on grammar and cognition in Pirahã: another look at the design features of human language.' http://lings.ln.man.ac.uk.

Everett D (2004b). 'Pirahã, a Native American language of northwest Brazil.' http://web.archive.org.

Gal S & Irvine J T (1995). 'The boundaries of languages and disciplines: how ideologies construct differences.' *Social Research 62(4),* 967–1001.

Golla V (2000). 'Language history and communicative strategies in aboriginal California and Oregon.' In Miyaoka O & Oshima M (eds.) *Languages of the north Pacific rim 5.* Suita, Japan: Faculty of Informatics, Osaka Gakuin University.
Hill J H (2000). 'Languages on the land.' In Terrell J (ed.) *Language, archaeology, and history.* Westport, CT: Bergin & Garvey.
Hill J H & Hill K C (1997). 'Culture influencing language: plurals of Hopi kin terms in comparative Uto-Aztecan perspective.' *Journal of Linguistic Anthropology 7,* 166–180.
Hunn E (2001). 'Prospects for the persistence of "endemic" cultural systems of traditional environmental knowledge: a Zapotec example.' In Maffi L (ed.) *On biocultural diversity: linking language, knowledge, and the environment.* Washington/London: Smithsonian Institution Press.
Hymes D H (1968). 'Functions of speech: an evolutionary approach.' In Cohen Y (ed.) *Man in adaptation: the biosocial background.* Chicago: Aldine.
Inoue M (2004). 'What does language remember? Indexical inversion and the naturalized history of Japanese women.' *Journal of Linguistic Anthropology 14,* 39–58.
Kari J (1989). 'Some linguistic insights into Dena'ina prehistory.' In Cook E-D & Rice K D (eds.) *Athapaskan linguistics: current perspectives on a language family.* Berlin/New York: Mouton de Gruyter.
Kroskrity P (1993). *Language, history, and identity: ethnolinguistic studies of the Arizona Tewa.* Tucson: University of Arizona Press.
MacLaury R E (1991). 'Social and cognitive motivations of change: measuring variability in color semantics.' *Language 67,* 34–62.
McConvell P (2001). 'Language shift and language spread among hunter-gatherers.' In Panter-Brick C, Rowley-Conwy P & Layton R (eds.) *Hunter-gatherers: cultural and biological perspectives.* Cambridge: Cambridge University Press.
Miller W R (1970). 'Western Shoshoni dialects.' In Swanson E H Jr (ed.) *Languages and cultures of North America.* Pocatello: Idaho State University Press.
Miller W R (1971). 'The death of language or serendipity among the Shoshoni.' *Anthropological Linguistics 13(3),* 114–120.
Milroy J & Milroy L (1985). 'Linguistic change, social network and speaker innovation.' *Journal of Linguistics 21,* 339–384.
Mufwene S (2001). *The ecology of language evolution.* Cambridge: Cambridge University Press.
Nichols J (1992). *Linguistic diversity in space and time.* Chicago: University of Chicago Press.
Renfrew C & Bellwood P (eds.) (2002). *Examining the farming/language dispersal hypothesis.* Cambridge: MacDonald Institute for Archaeological Research.
Silverstein M (1976). 'Shifters, linguistic categories, and cultural description.' In Basso K & Selby H (eds.) *Meaning in anthropology.* Albuquerque: University of New Mexico Press.
Silverstein M (1979). 'Language structure and linguistic ideology.' In Clyne P R, Hanks W F & Hofbauer C L (eds.) *The elements: a parasession on linguistic units and levels.* Chicago: Chicago Linguistic Society.
Silverstein M (1996). 'Indexical order and the dialectics of sociolinguistic life.' In Ide R, Parker R & Sunaoshi Y (eds.) *SALSA III: Proceedings of the Third Annual Symposium about Language and Society – Austin.* Austin: Department of Linguistics, University of Texas at Austin.
Thurston W F (1987). *Processes of change in the languages of northwestern New Britain.* Canberra: Pacific Linguistics.
Thurston W F (1989). 'How exoteric languages build a lexicon: esoterogeny in west New Britain.' In Harlow R & Hooper R (eds.) *VICAL 1, Oceanic Languages: Papers from the Fifth International Conference on Austronesian Linguistics.* Auckland: Linguistic Society of New Zealand.
Thurston W F (1992). 'Sociolinguistic typology and other factors effecting change in northwestern New Britain.' In Dutton T (ed.) *Culture change, language change: some case studies from Melanesia.* Canberra: Pacific Linguistics.
Urban G P (2001). *Metaculture: how culture moves through the world.* Minneapolis: University of Minnesota Press.
Whorf B L ([1941] 1956). 'The relation of habitual thought and behavior to language.' In Carroll J (ed.) *Language, thought, and reality: selected writings of Benjamin Lee Whorf.* Cambridge, MA: MIT Press.

Language Change and Language Contact

S Thomason, University of Michigan, Ann Arbor, MI, USA

Mainstream historical linguistics came rather late to the knowledge that language contact can, and often does, lead directly to structural linguistic changes. Leading figures expressed a firm belief in severe constraints on structural interference; Meillet, for instance, believed that grammatical loans can only occur when the source and receiving systems are very similar (Meillet, 1921: 87), and Jakobson argued that "a language accepts foreign structural elements only when they correspond to its own tendencies of development" (Jakobson, 1962 [1938]: 241). These beliefs probably arose ultimately from a conviction that the Comparative Method would be threatened

Golla V (2000). 'Language history and communicative strategies in aboriginal California and Oregon.' In Miyaoka O & Oshima M (eds.) *Languages of the north Pacific rim 5*. Suita, Japan: Faculty of Informatics, Osaka Gakuin University.
Hill J H (2000). 'Languages on the land.' In Terrell J (ed.) *Language, archaeology, and history.* Westport, CT: Bergin & Garvey.
Hill J H & Hill K C (1997). 'Culture influencing language: plurals of Hopi kin terms in comparative Uto-Aztecan perspective.' *Journal of Linguistic Anthropology 7,* 166–180.
Hunn E (2001). 'Prospects for the persistence of "endemic" cultural systems of traditional environmental knowledge: a Zapotec example.' In Maffi L (ed.) *On biocultural diversity: linking language, knowledge, and the environment.* Washington/London: Smithsonian Institution Press.
Hymes D H (1968). 'Functions of speech: an evolutionary approach.' In Cohen Y (ed.) *Man in adaptation: the biosocial background.* Chicago: Aldine.
Inoue M (2004). 'What does language remember? Indexical inversion and the naturalized history of Japanese women.' *Journal of Linguistic Anthropology 14,* 39–58.
Kari J (1989). 'Some linguistic insights into Dena'ina prehistory.' In Cook E-D & Rice K D (eds.) *Athapaskan linguistics: current perspectives on a language family.* Berlin/New York: Mouton de Gruyter.
Kroskrity P (1993). *Language, history, and identity: ethnolinguistic studies of the Arizona Tewa.* Tucson: University of Arizona Press.
MacLaury R E (1991). 'Social and cognitive motivations of change: measuring variability in color semantics.' *Language 67,* 34–62.
McConvell P (2001). 'Language shift and language spread among hunter-gatherers.' In Panter-Brick C, Rowley-Conwy P & Layton R (eds.) *Hunter-gatherers: cultural and biological perspectives.* Cambridge: Cambridge University Press.
Miller W R (1970). 'Western Shoshoni dialects.' In Swanson E H Jr (ed.) *Languages and cultures of North America.* Pocatello: Idaho State University Press.
Miller W R (1971). 'The death of language or serendipity among the Shoshoni.' *Anthropological Linguistics 13(3),* 114–120.
Milroy J & Milroy L (1985). 'Linguistic change, social network and speaker innovation.' *Journal of Linguistics 21,* 339–384.
Mufwene S (2001). *The ecology of language evolution.* Cambridge: Cambridge University Press.
Nichols J (1992). *Linguistic diversity in space and time.* Chicago: University of Chicago Press.
Renfrew C & Bellwood P (eds.) (2002). *Examining the farming/language dispersal hypothesis.* Cambridge: MacDonald Institute for Archaeological Research.
Silverstein M (1976). 'Shifters, linguistic categories, and cultural description.' In Basso K & Selby H (eds.) *Meaning in anthropology.* Albuquerque: University of New Mexico Press.
Silverstein M (1979). 'Language structure and linguistic ideology.' In Clyne P R, Hanks W F & Hofbauer C L (eds.) *The elements: a parasession on linguistic units and levels.* Chicago: Chicago Linguistic Society.
Silverstein M (1996). 'Indexical order and the dialectics of sociolinguistic life.' In Ide R, Parker R & Sunaoshi Y (eds.) *SALSA III: Proceedings of the Third Annual Symposium about Language and Society – Austin.* Austin: Department of Linguistics, University of Texas at Austin.
Thurston W F (1987). *Processes of change in the languages of northwestern New Britain.* Canberra: Pacific Linguistics.
Thurston W F (1989). 'How exoteric languages build a lexicon: esoterogeny in west New Britain.' In Harlow R & Hooper R (eds.) *VICAL 1, Oceanic Languages: Papers from the Fifth International Conference on Austronesian Linguistics.* Auckland: Linguistic Society of New Zealand.
Thurston W F (1992). 'Sociolinguistic typology and other factors effecting change in northwestern New Britain.' In Dutton T (ed.) *Culture change, language change: some case studies from Melanesia.* Canberra: Pacific Linguistics.
Urban G P (2001). *Metaculture: how culture moves through the world.* Minneapolis: University of Minnesota Press.
Whorf B L ([1941] 1956). 'The relation of habitual thought and behavior to language.' In Carroll J (ed.) *Language, thought, and reality: selected writings of Benjamin Lee Whorf.* Cambridge, MA: MIT Press.

Language Change and Language Contact

S Thomason, University of Michigan, Ann Arbor, MI, USA

Mainstream historical linguistics came rather late to the knowledge that language contact can, and often does, lead directly to structural linguistic changes. Leading figures expressed a firm belief in severe constraints on structural interference; Meillet, for instance, believed that grammatical loans can only occur when the source and receiving systems are very similar (Meillet, 1921: 87), and Jakobson argued that "a language accepts foreign structural elements only when they correspond to its own tendencies of development" (Jakobson, 1962 [1938]: 241). These beliefs probably arose ultimately from a conviction that the Comparative Method would be threatened

by the existence of extensive structural diffusion. It is no accident that Hugo Schuchardt, a major critic of what he saw as the Neogrammarians' intellectual rigidity in ignoring foreign interference, was the founder of pidgin/creole studies, because mixed languages like pidgins and creoles challenge the universality of the tidy family tree (see Thomason and Kaufman, 1988). In general, however, historical linguists have long tended to argue that a contact explanation for a change should be proposed only when all attempts to find a language-internal explanation have failed. Research over the past half century, starting with Weinreich's classic 1953 book *Languages in contact*, has brought about a recognition of the importance of language contact for explanations of many linguistic changes. In recent years, the trickle of monographs and even textbooks on language contact has become a flood (two of the most recent important treatments are Winford, 2003 and Clyne, 2003).

The following sections will address the general problem of how to determine, after the fact, that contact-induced change has occurred. The first section focuses on the clearest cases, those for which no one would deny the effects of contact. In the second section, less obvious cases will be considered, with conclusions on methods for establishing the existence of prior contact-induced changes. Contact-induced changes that do not involve actual interference will be considered in the third section. The concluding section will offer some comments on the prospects for predicting what kinds of contact-induced changes will occur when.

Before turning to a detailed discussion of the issues that arise in this area, we need a definition of contact-induced language change: contact is a cause of "any linguistic change that would have been less likely to occur outside a particular contact situation" (Thomason, 2001: 62). This definition covers all changes in which foreign material is transferred from one language to another, i.e., all cases of linguistic interference, but it also includes some less obvious types of change. First, some changes that occur in some cases of slow language death fall into the category of attrition – loss of linguistic material – but do not make the dying language more similar to the language that is replacing it; these are nevertheless contact-induced changes by my definition. Second, intentional linguistic changes, for instance in cases where a speech community deliberately distances its language from neighboring languages, are contact-induced but do not involve diffusion. And third, some changes occur as an indirect result of interference, typically when a borrowed morpheme sets off a chain reaction that has a snowballing effect on the receiving language's structure. Finally, the definition does not exclude contribution from internal pattern pressures in a relevant change: multiple causation is always a possibility, whether the causes are all internal or a mixture of internal and contact factors.

Contact-Induced Change: Uncontroversial Cases

A great many instances of contact-induced change will of course be accepted by all historical linguists without hesitation. The vast majority of these are changes in which morphemes are transferred from one language to another – namely, lexical borrowing and the borrowing of grammatical morphemes, both particles and affixes. No one questions whether English words such as *shah* (from Persian) and *chaise* (from French) are native or of foreign origin; similarly, no one doubts the foreign origin of English suffixes such as *-able* in words like *readable*, which has a native Germanic root, or of particles such as the Huastec (Mayan) conjunction *pero* 'but' (from Spanish). Borrowings like these arguably have no immediate structural consequences in themselves, because they do not alter native morphosyntax.

Other borrowings, however, have a clear structural effect on the receiving language. So, for example, the Indic language Bengali has borrowed the suffixes *-ra*: 'human plural' and *-gulo:, guli* 'nonhuman plural,' together with the noun-class categories they express, from Dravidian (Andronov, 1964: 124). The development of a vowel harmony rule in some dialects of Asia Minor Greek surely started with Greek suffixes added to Turkish loanwords and then spread to native Greek vocabulary. Dawkins, who reports this phenomenon, comments that "the fullness with which vowel harmony is observed clearly depends on how far the individual speaker is accustomed to talk Turkish and has the Turkish ear for these distinctions" (Dawkins, 1916: 68).

Examples of both types, with and without immediate structural effects on the receiving language, can be multiplied endlessly. The most extreme of the clear cases are mixed languages – the small set of bilingual mixed languages and the much larger set of pidgin and creole languages. For all the languages in these sets that arose abruptly (including most known bilingual mixed languages and most pidgins, but apparently fewer creoles), there is a question as to whether one can reasonably talk about language change at all: processes of language creation closely resemble processes of language change in less dramatic situations, but the end result is a new language, not a changed later form of a single language. This is not the place to try to answer this question, but mixed languages are worth mentioning here because they

enjoy a privileged position among the linguistic outcomes of contact. One might imagine (with difficulty) an argument that Asia Minor Greek just happened to develop vowel harmony, and that the development had nothing to do with Turkish influence, in spite of all the Turkish loanwords that display vowel harmony. But unlike other products of language contact, mixed languages declare their mixed origin as soon as any attempt is made to include them in an application of the Comparative Method, because they always display a striking discrepancy between one (or more) grammatical subsystems and the rest of the language. They are therefore the easiest products of language contact to detect retrospectively, provided that at least some of their source languages are still available for comparison. Pidgins and creoles usually (though not always) draw their lexicon primarily from a single language, but the grammar as a whole comes neither from that language nor from any other single language. Bilingual mixed languages differ in their make-up: Michif, a French-Cree mixture spoken mainly in Manitoba and North Dakota, has noun phrases (lexicon, phonology, morphology, and syntax) from French and the rest of the language (lexicon, phonology, morphology, and syntax) from Cree; Mednyj Aleut, once spoken on one of the Commander Islands off the far eastern coast of Russia, has Russian finite verb morphology but other grammatical components (including most of the lexicon) from Aleut; Media Lengua (Ecuador) has Spanish lexicon and Quechua grammar; and so forth.

For all these cases of contact-induced change and language creation under contact conditions, no special methodology is needed to identify the effects of language contact, always assuming that the source languages, or their relatives, are known. In other contact situations, however, the picture is less clear, and the need for rigorous methods for establishing the existence of contact-induced change is evident.

Less Obvious Contact-Induced Changes

In the borrowing situations just described, contact-induced change is signaled by the presence of foreign material incorporated into the receiving language and is therefore uncontroversial. In other cases, structural features are transferred without morphemes. In fact, this is very common in three kinds of contact situations. First, when the source language and the receiving language are very closely related, with essentially no typological distance separating them and largely shared lexicon and structural features, structural diffusion from one to the other often happens without any morphemic exchange. This is perhaps most obvious in cases of dialect borrowing, as for instance when a Standard Serbo-Croatian pattern of syncretism in plural oblique cases (genitive vs. dative/instrumental/locative) replaced the naive syncretic pattern of the Hvar dialect (genitive/locative vs. dative/instrumental), though Hvar speakers retained their own case suffixes, which differ from those of the standard dialect (Hraste, 1935: 17–25).

The second type of contact situation in which there is very often structural interference without morpheme transfer is in cases of language shift (typically substratum interference). In borrowing situations of the sorts discussed above, where morpheme transfer accompanies structural interference, nonbasic vocabulary is the earliest and the predominant linguistic effect; structural interference is likely only with increasingly intense contact involving widespread bilingualism. In sharp contrast, phonological and syntactic features predominate in shift-induced interference. The reason is that imperfect learning of a target language plays a (or the) major role in shift-induced interference and, as is typical in second-language learning, vocabulary is learned first (see Thomason and Kaufman, 1988: Chap. 3 and Thomason, 2001: 66–76 for discussion of this distinction). (Inflectional morphology tends to be least common in both types of interference, for reasons having to do with its typically tight-knit integration into the grammatical framework of a language.)

The third type of interference without morpheme transfer is (at least as far as is currently known) less common than shift-induced interference and much less common that borrowing between closely related languages. This is the convergence of structure through mutual or one-way borrowing in a highly organized intense bilingual or multilingual contact situation in which each speaker group is concerned to preserve its own heritage language – sometimes, though by no means always, with a reluctance to engage in any lexical borrowing. A small number of cases have been reported in the literature for which word-by-word or even morpheme-by-morpheme intertranslatability is claimed for the languages; unfortunately, however, few thorough detailed case studies are available for any of these situations (but see Aikhenvald, 2003 for a report on an Amazonian contact situation).

The most famous case is the village of Kupwar in India, where the Indic languages Marathi and Urdu and the Dravidian language Kannaḍa are said to have converged drastically in their structures (Gumperz and Wilson, 1971; see also Thomason and Kaufman, 1988: 86–88 for discussion). It is unclear just how drastic the convergence was in this case, given that Urdu and Marathi are related and that Dravidian and

Indic languages have been in such close contact for so long that the typological distance between the two groups has been lessened by much contact-induced change. Even so, the typological match in the Kupwar languages is striking. Of the sixteen features discussed by Gumperz and Wilson, all of them morphosyntactic, Kupwar Urdu has converged toward Kupwar Kannaḍa and/or Marathi in twelve, Kupwar Kannaḍa has converged toward Marathi in five or six, and Kupwar Marathi has converged toward Kannaḍa in four features. Among the features are gender categories, an inclusive/exclusive 'we' distinction, specific subject-verb agreement rules in several constructions, and yes/no question marking. Only one of the changes, the adoption by Kannaḍa of an Indic subordinator *ki* together with some (not all) of its syntactic features, involves a borrowed morpheme.

The difficulty with interference that does not include transferred morphemes is that those who consider contact-induced change an exotic and inherently unlikely explanation for structural change find it easier to maintain their skepticism about contact explanations in these cases. The Ethiopian Highlands, the best-known linguistic area in Africa, comprises Ethiopic Semitic languages in contact with Cushitic languages (see especially the ground-breaking work of Leslau, 1945, 1952). Semitic and Cushitic, in spite of their genetic link as coordinate branches of the Afro-Asiatic language family, are typologically quite different in many respects. Ethiopic Semitic shares a sizable number of structural features with Cushitic, in contradistinction to the rest of Semitic, and Leslau attributes these sharings to the influence of Cushitic speakers who shifted to Semitic. Although Ethiopic Semitic also has numerous Cushitic loanwords, the most striking morphosyntactic interference features employ native Semitic morphemes. The most-affected Ethiopic Semitic dialects share with Cushitic otherwise un-Semitic features such as labialized velar consonants, a distinct future tense, new morphological means of expressing the causative, negative perfect, and frequentative in verbs, noun plurals formed by reduplication of the root-final consonant, subordinate clauses expressed by nonfinite gerund constructions, and several linked word-order features – SOV sentential word order, auxiliary following the main verb, relative clauses preceding the head noun, and postpositions instead of prepositions.

The problem of establishing the existence of contact-induced change is especially acute when the changes could easily have happened through internally motivated change. When the proposed interference features are universally marked, as with the Ethiopic Semitic labialized velars, or when they belong to the most tightly organized closed grammatical subsystems, such as the causative, negative perfect, and frequentative verbal formations and the reduplicative plurals of Ethiopic Semitic, the chances of independent internal innovation are less; but when the possibilities are limited, as with basic sentential word order and linked word-order features (postpositions vs. prepositions and the like), internal motivations could easily have produced the same outcome as contact. Adopting a rigorous methodology will help in such cases. It will not always provide a definitive answer to the contact and/or-internal-causation question – many things are unknowable in historical linguistics, as in other historical sciences – but it will at least show whether a definitive answer is possible (see Thomason, 2001: 93–94).

First, it is necessary to look at the whole picture, not just a single piece of the puzzle. If one syntactic feature has diffused from language A to language B, other syntactic features and some phonological features as well will almost certainly have done so: a contact situation intense enough to permit structural interference at one structure point also permits structural interference at other points. Second, we need to identify a source language for the set of features we wish to attribute to language contact. Appeals to unidentified substrata are justifiably viewed with disapproval by historical linguists, because they do not advance our knowledge. In cases of language shift, of course, the whole point is that the shifting group's language often disappears; but if another subgroup of the speech community maintains the same language, or if there are closely related languages that can be investigated, this problem can be overcome. Third, we need to prove that language B has changed (no change, no evidence for interference); and fourth, we need to prove that the proposed interference features are old in language A (if they are innovative in A, then A is hardly likely to be the source of the features in B). A complete investigation will also consider the possibility of multiple causation and search for potential internal motivations as well.

Attempts to apply these criteria will not always be successful. The languages of the famous Pacific Northwest linguistic area of the northwestern United States and neighboring British Columbia are a case in point. They share many striking features (some of them highly marked universally): labialized dorsal consonants, velars vs. uvulars, ejectives, verb-initial sentential word order, sentence-initial negation, many suffixes but relatively few (or no) prefixes, a weak noun vs. verb distinction, the use of reduplication to mark a distributive plural, numeral classifiers, large sets of lexical suffixes (with meanings like 'long object,' 'belly,' and 'face/fire'), etc. But since all

these features must be reconstructed for all three proto-languages of the core families in the area – Salishan, Wakashan, and Chemakuan – evidence for the original source(s) of the shared features is lacking, and in fact there is no evidence for **any** diffusion involving these features. This does not of course mean that the Pacific Northwest is not in fact a linguistic area; it does mean that we cannot prove it at present, for the area as a whole. It seems extremely unlikely that all these features just happened to arise independently in the three families. But distant genetic relationship is a possibility, though solid evidence for that is also lacking, so that ancient areal diffusion is still the most likely explanation for the typological congruence.

Although sticking to a rigorous methodology sometimes precludes a firm conclusion about the linguistic results of a historical contact situation, it will help to evaluate poorly supported proposals of contact-induced change.

Contact-Induced Changes that Do Not Involve Direct Interference

As noted above, several change processes that fall under my definition of contact-induced change do not result in the diffusion of any lexical or structural features from one language to another. The first of these is connected with language death. Not all dying languages undergo the kinds of changes that are the focus of most of the literature on language death (see for example the papers in Dorian, 1989); some languages vanish without undergoing any significant changes at all, especially in rapid shift situations. But languages that lose speakers gradually very often also lose their lexical and structural resources gradually, through two quite distinct processes. First, their speakers replace native lexicon and structure through borrowing from the dominant language to which they are shifting; and second, speakers abandon native lexicon and structure without any replacements – a process known as attrition. These two processes are often responsible for the same changes, specifically when a borrowed structure is simpler, or less marked universally, than the structure it replaces in the receiving language. A detailed study of one case, a Hungarian-American community near Pittsburgh, Pennsylvania, provided very few structural examples of attrition alone (that is, examples in which the lost structures did not bring Hungarian closer to English structure); more examples of borrowing alone (closer to English but no simpler than the replaced Hungarian structures); and still more examples, though not many more, of attrition plus borrowing (closer to English and also simpler than the replaced Hungarian structures) (Fenyvesi, 1995). So, for instance, this variety of American Hungarian has reduced the inherited system of preverbs, a change that is possibly due to attrition alone but is probably best analyzed as a combination of attrition and borrowing (since English has no comparable preverb system).

Another class of contact-induced changes without interference is to be found in certain instances of intentional change. These fall into two different (though related) categories: cases in which a group deliberately alters its language to make it more different from the language(s) of neighboring groups, as a means of underlining in-group identity, and cases where a group invents a secret language in order to keep conversations from being understood by hostile outsiders. Examples of the former often involve massive distortion of the lexicon. Speakers of Lambayeque Quechua, for instance, metathesized certain vowel-consonant sequences in words because they felt that their language was too similar to their neighbors' language (Dwight Shaver, as reported by David Weber, personal communication, 1999). And in an instance of the latter type, the Lōṛīs of what is now Pakistan are reported to have invented a secret jargon, Mōkkī, by distorting their lexicon (including loanwords from Baluchi- and Brahui-speaking neighbors) by metathesis and by adding prefixes and suffixes (Bray, 1913: 139–140). This example might have only passing interest as a curiosity rather more elaborate than Pig Latin and other language games, except that Bray reports that Mōkkī is "acquired naturally and as a matter of course by Lōṛī children; . . . it is becoming a language for the home-circle . . ." (Bray, 1913: 139–140).

A somewhat different sort of language distortion for privacy purposes is reported in a number of cases from several different continents, cases in which groups deliberately simplify, or pidginize, their native language in order (as they assert) to keep outsiders from learning it. The Delaware Indians did this with their language, and although some foreigners (traders and missionaries) did not notice the distortion, others did: the Dutch missionary Michaëlius commented in 1628 that they "rather design to conceal their language from us than to properly communicate it, except in things that happen in daily trade; saying that it is sufficient for us to understand them in that" (Jameson, 1967 [1909]: 128). Similarly, a field-worker in Ethiopia reported that, after spending 7 months working diligently to learn the Hamer language, she realized that what she had actually learned was "a kind of 'Pidgin Hamer' which is used only for and by policemen, traders, and non-Hamer settlers"

(Lydall, 1976: 128). And an early missionary to the Motus in New Guinea discovered to his chagrin that – as he was informed by his young son, who spent much time with the village boys – the language he had struggled to learn was 'a simplified form' of the Motu language, because the Motus were reluctant to teach him their 'true' language (Dutton and Brown, 1977: 760–761, citing Chatterton, 1970: 95–96; the missionary was W. G. Lawes, who arrived in Port Moresby in 1874).

The third category of contact-induced change without interference requires a modifier: changes in this category do result from linguistic interference, but the structural effects are delayed, not immediate. That is, the borrowed element (these changes are typical of borrowing situations, not of shift-induced interference) does not in itself constitute structural change, but its presence triggers a series of changes that do eventually bring about significant structural change, in a kind of snowball effect. The borrowing of conjunctions into languages in which subordination was traditionally expressed by nonfinite clauses often leads to an innovative system with finite subordinate clauses. A comparable example is the change in Siberian Yupik (Eskimo) from a suffixal NP coordination pattern, with a comitative suffix added to each noun, to a coordination pattern in which the two NPs are linked by a conjunction, 'and,' borrowed from Chukchi. Menovščikov (1969: 124–130) observed three different coordination patterns that apparently marked a transition from the old to the new pattern: the old construction alone, with the comitative suffixes; a mixed construction with the suffixes and the conjunction 'and'; and an innovative construction with the conjunction alone. The development of this innovative end-stage construction was not itself due to borrowing from Chukchi, but the process that led to the current situation was set in motion by the adoption of the Chukchi conjunction.

Caution is required in the analysis of chain-reaction changes in this category: as Mark Hale has observed (personal communication 2002), in principle there is no limit to how far back one could reach into a language's history to find a borrowed element and then argue that it motivated a late change in a chain reaction. Few historical linguists would wish to argue for a Greek source for a change in (say) a modern Portuguese dialect, even if a morpheme involved in the change can ultimately be traced back to a Greek loanword in early Latin. The most convincing cases involve relatively shallow time depths, with chain reactions stretching back a few hundred years at most.

Conclusion: Can We Predict What Kinds of Contact-Induced Changes Will Occur When?

Historical linguists tend to be pessimists when it comes to predicting change. Nevertheless, we are always on the lookout for linguistic and social predictors for change, and especially for constraints on what changes are possible. In language contact, the safest prediction is completely trivial and therefore uninteresting: no language contact, no contact-induced change. More substantive predictions and constraints are hard to find. A common assumption is that contact-induced change is inevitable under conditions of intense contact, i.e., when there is very widespread bilingualism in a speech community. But even this commonsense position is not a safe assumption, because there are well-documented cases of bilingual speech communities whose members do not engage in lexical borrowing, and in some of these situations there is also no solid evidence of structural interference. Montana Salish, for example, does not in general borrow words from English (or French, the other European language that has played a prominent role in its contact history), and it has acquired no structural features from English (or French) either. There has certainly been a great deal of cultural assimilation to the dominant anglo culture over the past 150 years or so, but the words that are used to designate new items are created from native material; the word for 'automobile', to give just one of many examples, is *p'ip'úyšn*, which literally means 'it has wrinkled feet'. Moreover, this appears to be an areal feature of the U.S. Northwest, because Nez Perce and other languages of the region also tend not to borrow words from English or French. A few current Montana Salish speakers occasionally produce sentences that show some English influence, e.g., with SVO word order. But although such sentences are fully grammatical in Montana Salish, they are stylistically marked. They occur very infrequently in naturally occurring data, such as stories, and are common only for a few speakers and only under sentence-elicitation conditions. It seems risky, therefore, to suppose that contact-induced change of any sort is inevitable in intense contact.

Strenuous efforts to identify unborrowable linguistic features have also proved fruitless: all the constraints that have been proposed so far can be, and indeed have been, proved to be leaky. But even if we cannot talk safely about possibilities, we can still consider probabilities. There is strong evidence to support the claim that social factors outweigh linguistic factors in predicting the linguistic results of

contact. It is unfortunate, therefore, that our understanding of the crucial social factors remains fragmentary. The distinction outlined above, between borrowing (adopting a feature from one language into another, where the borrowers are fluent in both languages) and shift-induced interference (where imperfect learning of a target language is crucial), correlates robustly with linguistic effects: nonbasic vocabulary first and most in borrowing, with structure and basic vocabulary borrowed later if at all; phonology and syntax most prominent in shift-induced interference, with lexical transfer lagging behind or absent altogether. One implication of this distinction is that it permits an educated guess about the type of contact that was responsible for contact-induced changes in a long-vanished contact situation: if we can establish a number of structural interference features but no lexical interference, we can be almost certain that the process was shift-induced interference. (A rare exception might be borrowing by bilinguals into a language whose speakers have an explicit or implicit ban on borrowing vocabulary.)

Other social factors, most notably intensity of contact, have been mentioned above and in the literature, but the concept is distressingly vague. In borrowing situations, it has to do in part with levels of bilingualism in the speech community. In shift situations, the relative sizes of the two speech communities and the speed of the shift contribute to the level of intensity: the more shifting speakers there are relative to speakers of the target language, and the more rapid the shift, the more interference features one can expect in the target language.

Linguistic factors, though important, are easily overridden by social factors. For instance, typological distance between source language and receiving language affects the likelihood that structure will be borrowed – the more similar the systems are to begin with, the easier it is for a feature to diffuse from one to the other – but with intense enough contact, **any** feature can be transferred from any language to any other language, no matter how different the two languages are typologically.

Considerations of universal markedness also play a role, but mainly in shift situations. In borrowing situations, where the receiving-language speakers are fluent bilinguals, knowledge of the relevant structures can safely be assumed, so there is no knowledge-based barrier to borrowing. In shift situations, by contrast, markedness plays a role in the learning process: highly marked target-language features are less likely to be learned by members of the shifting group. On the other hand, when shifting speakers carry over features of their heritage language into their version of the target language, these features may be marked or not; but the unmarked first-language features are less likely than the marked ones to be adopted by original target-language speakers if the two speech communities coalesce into a single integrated speech community. That is, in such cases the original target-language speakers will adopt some, but not all, of the innovative features in the shifting group's variety of the target language; and the less marked features, being easier to learn, are more likely to survive in the eventual integrated target language.

Markedness itself is affected by various specific linguistic factors. To give just one example, one common-sense and valid proposal is that sharpness of boundaries makes affixes easier to borrow (Heath, 1978: 105; Comrie, 1981: 202–203). This leads to a prediction that affixes from flexional languages, and also from agglutinative languages where morpheme boundaries are obscured by morphophonemic alternations, will be less often borrowed than affixes from agglutinative languages with transparent transitions from one morpheme to the next in a string of stem and affixes.

Various authors have proposed borrowing scales, correlating expected (or permissible) interference features with increasing intensity of contact: the more intense the contact, the more kinds of features can be borrowed (see e.g., Moravcsik, 1978 and Thomason and Kaufman, 1988: 73ff). These scales are roughly predictive, in that they embody a claim that, for instance, stance, basic vocabulary will not be borrowed before nonbasic vocabulary, and inflectional morphology will not be borrowed before derivational morphology. Occasional exceptions to these predictions sometimes turn up, for complex social and linguistic reasons, so no borrowing scale is completely accurate. Perhaps the most common reason for out-of-scale borrowings is close typological congruence between the source and receiving languages, which can facilitate types of borrowing under less intense contact conditions than the scale predicts. Finally, different authors' borrowing scales often make different predictions; Romaine predicts, for example, that morphology will be borrowed before syntax (Romaine, 1995: 64), while Thomason and Kaufman (1988) predict the opposite borrowing order.

The conclusion, then, is that rough predictions can be made about the likelihood of different kinds and degrees of both borrowing and shift-induced interference, under varying social and linguistic conditions. Many of these predictions are for the most part borne out by the available empirical evidence, and some are strongly supported. But the possibility for exceptions

always exists, because no aspect of language change is, in the end, completely predictable.

Bibliography

Aikhenvald A (2003). *Language contact in Amazonia.* Oxford: Oxford University Press.

Andronov M (1964). 'On the typological similarity of New Indo-Aryan and Dravidian.' *Indian Linguistics 25,* 119–126 (Baburam Saksena Felicitation Volume).

Bray D & de S (1913). *Census of India, vol. IV: Baluchistan.* Calcutta: Superintendent Government Printing, India.

Chatterton P (1970). 'The origin and development of Police Motu.' *Kivung 3,* 95–98.

Clyne M (2003). *Dynamics of language contact: English and immigrant languages.* Cambridge: Cambridge University Press.

Comrie B (1981). *Language universals and linguistic typology.* Chicago: University of Chicago Press.

Dawkins R M (1916). *Modern Greek in Asia Minor: a study of the dialects of Sílli, Cappadocia and Phárasa with grammars, texts, translations, and glossary.* Cambridge: Cambridge University Press.

Dorian N C (ed.) (1989). *Investigating obsolescence: studies in language contraction and death.* Cambridge: Cambridge University Press.

Dutton T & Brown H A (1977). 'Hiri Motu: the language itself.' In Wurm S A (ed.) *New Guinea area languages and language study,* vol. 3. Canberra: Australian National University. 759–793.

Fenyvesi A (1995). 'Language contact and language death in an immigrant language: the case of Hungarian.' *University of Pittsburgh Working Papers in Linguistics 3,* 1–117.

Gumperz J & Wilson R (1971). 'Convergence and creolization: a case from the Indo-Aryan/Dravidian border.' In Hymes D (ed.) *Pidginization and creolization of languages.* Cambridge: Cambridge University Press. 151–167.

Heath J (1978). *Linguistic diffusion in Arnhem Land.* Canberra: Australian Institute of Aboriginal Studies.

Hraste M (1935). 'Cakavski Dijalekat Ostrva Hvara.' *Južnoslovenski Filolog 14,* 1–59.

Jakobson R (1962). 'Sur la théorie des affinités phonologiques entre des langues.' In Jakobson R (ed.) *Selected Writings,* vol. 1. 234–246. (reprinted from *Actes du Quatrième Congrès International de Linguistes*, Copenhagen: Munksgaard, 1938. 48–59.) The Hague: Mouton.

Jameson J F (ed.) (1967 [1909]). *Narratives of New Netherland 1609–1664.* New York: Charles Scribner's Sons; reprinted, New York: Barnes & Noble, 1967.

Leslau W (1945). 'The influence of Cushitic on the Semitic languages of Ethiopia: a problem of substratum.' *Word 1,* 59–82.

Leslau W (1952). 'The influence of Sidamo on the Ethiopic languages of the Gurage.' *Language 28,* 63–81.

Lydall J (1976). 'Hamer.' In Lionel Bender M (ed.) *The non-Semitic languages of Ethiopia.* East Lansing, MI: African Studies Center. 393–438.

Meillet A (1921). 'Le problème de la parenté des langues.' In Meillet A (ed.) *Linguistique historique et linguistique générale.* Paris: Champion. 76–101.

Menovšikov G A (1969). 'O nekotoryx social 'nyx aspektax evoljucii jazyka.' In *Voprosy social 'noj lingvistiki.* Leningrad: Nauka. 110–134.

Moravcsik E (1978). 'Language contact.' In Greenberg J H (ed.) *Universals of human language, vol. 1: Method & Theory.* Stanford, CA: Stanford University Press. 93–122.

Romaine S (1995). *Bilingualism* (2nd edn.). Oxford: Blackwell.

Thomason S G (2001). *Language contact: an introduction.* Edinburgh: Edinburgh University Press/Washington DC: Georgetown University Press.

Thomason S G & Kaufman T (1988). *Language contact, creolization, and genetic linguistics.* Berkeley: University of California Press.

Weinreich U (1968 [1953]). *Language in contact.* The Hague: Mouton.

Winford D (2003). *An introduction to contact linguistics.* Oxford: Blackwell.

Language Change, Intentional

S Thomason, University of Michigan, Ann Arbor, MI, USA

The traditional view of linguistic change is that it is inexorable, beyond the control of the language's speakers either individually or in groups. This view has been expressed repeatedly over the past 150 years and across deep theoretical divides. Here are a few typical examples:

- It is 'not in the power of Man to produce or prevent changes in language' (Max Müller, *Lectures on the Science of Language*, 1861).
- 'Sometimes we can feel where the drift [of a language] is taking us even while we struggle against it' (Edward Sapir, *Language*, 1921: 156).
- Noam Chomsky and Morris Halle assume 'the adult's inability to modify his grammar except by the addition or elimination of a few rules' and assign the initiation of significant language change

pidgins and creoles
secret languages
social predicators of change
structural borrowing
typological distance

always exists, because no aspect of language change is, in the end, completely predictable.

Bibliography

Aikhenvald A (2003). *Language contact in Amazonia.* Oxford: Oxford University Press.

Andronov M (1964). 'On the typological similarity of New Indo-Aryan and Dravidian.' *Indian Linguistics 25,* 119–126 (Baburam Saksena Felicitation Volume).

Bray D & de S (1913). *Census of India, vol. IV: Baluchistan.* Calcutta: Superintendent Government Printing, India.

Chatterton P (1970). 'The origin and development of Police Motu.' *Kivung 3,* 95–98.

Clyne M (2003). *Dynamics of language contact: English and immigrant languages.* Cambridge: Cambridge University Press.

Comrie B (1981). *Language universals and linguistic typology.* Chicago: University of Chicago Press.

Dawkins R M (1916). *Modern Greek in Asia Minor: a study of the dialects of Sílli, Cappadocia and Phárasa with grammars, texts, translations, and glossary.* Cambridge: Cambridge University Press.

Dorian N C (ed.) (1989). *Investigating obsolescence: studies in language contraction and death.* Cambridge: Cambridge University Press.

Dutton T & Brown H A (1977). 'Hiri Motu: the language itself.' In Wurm S A (ed.) *New Guinea area languages and language study,* vol. 3. Canberra: Australian National University. 759–793.

Fenyvesi A (1995). 'Language contact and language death in an immigrant language: the case of Hungarian.' *University of Pittsburgh Working Papers in Linguistics 3,* 1–117.

Gumperz J & Wilson R (1971). 'Convergence and creolization: a case from the Indo-Aryan/Dravidian border.' In Hymes D (ed.) *Pidginization and creolization of languages.* Cambridge: Cambridge University Press. 151–167.

Heath J (1978). *Linguistic diffusion in Arnhem Land.* Canberra: Australian Institute of Aboriginal Studies.

Hraste M (1935). 'Cakavski Dijalekat Ostrva Hvara.' *Južnoslovenski Filolog 14,* 1–59.

Jakobson R (1962). 'Sur la théorie des affinités phonologiques entre des langues.' In Jakobson R (ed.) *Selected Writings,* vol. 1. 234–246. (reprinted from *Actes du Quatrième Congrès International de Linguistes*, Copenhagen: Munksgaard, 1938. 48–59.) The Hague: Mouton.

Jameson J F (ed.) (1967 [1909]). *Narratives of New Netherland 1609–1664.* New York: Charles Scribner's Sons; reprinted, New York: Barnes & Noble, 1967.

Leslau W (1945). 'The influence of Cushitic on the Semitic languages of Ethiopia: a problem of substratum.' *Word 1,* 59–82.

Leslau W (1952). 'The influence of Sidamo on the Ethiopic languages of the Gurage.' *Language 28,* 63–81.

Lydall J (1976). 'Hamer.' In Lionel Bender M (ed.) *The non-Semitic languages of Ethiopia.* East Lansing, MI: African Studies Center. 393–438.

Meillet A (1921). 'Le problème de la parenté des langues.' In Meillet A (ed.) *Linguistique historique et linguistique générale.* Paris: Champion. 76–101.

Menovšikov G A (1969). 'O nekotoryx social 'nyx aspektax evoljucii jazyka.' In *Voprosy social 'noj lingvistiki.* Leningrad: Nauka. 110–134.

Moravcsik E (1978). 'Language contact.' In Greenberg J H (ed.) *Universals of human language, vol. 1: Method & Theory.* Stanford, CA: Stanford University Press. 93–122.

Romaine S (1995). *Bilingualism* (2nd edn.). Oxford: Blackwell.

Thomason S G (2001). *Language contact: an introduction.* Edinburgh: Edinburgh University Press/Washington DC: Georgetown University Press.

Thomason S G & Kaufman T (1988). *Language contact, creolization, and genetic linguistics.* Berkeley: University of California Press.

Weinreich U (1968 [1953]). *Language in contact.* The Hague: Mouton.

Winford D (2003). *An introduction to contact linguistics.* Oxford: Blackwell.

Language Change, Intentional

S Thomason, University of Michigan, Ann Arbor, MI, USA

The traditional view of linguistic change is that it is inexorable, beyond the control of the language's speakers either individually or in groups. This view has been expressed repeatedly over the past 150 years and across deep theoretical divides. Here are a few typical examples:

- It is 'not in the power of Man to produce or prevent changes in language' (Max Müller, *Lectures on the Science of Language*, 1861).
- 'Sometimes we can feel where the drift [of a language] is taking us even while we struggle against it' (Edward Sapir, *Language*, 1921: 156).
- Noam Chomsky and Morris Halle assume 'the adult's inability to modify his grammar except by the addition or elimination of a few rules' and assign the initiation of significant language change

to children's first-language acquisition (*The Sound Pattern of English*, 1968: 251).

- The 'mistake [in functionalist explanatory strategies] is considering language change to be something that speakers "do" (in any real sense, for any reason), rather than something that happens to their languages . . .' (Lass, 1997: 370).

The regularity hypothesis of sound change, given its great importance for historical linguistic methodology, may be the main reason for the traditional insistence that speakers cannot initiate and implement deliberate structural changes in their languages, because its validity rests on the assumption that sound change proceeds independently of speakers' intentions (and of grammatical context). Presumably no historical linguist would wish to deny that the vast majority of structural linguistic changes fit this picture. Nevertheless, deliberate and conscious decisions taken by speakers can be shown to be responsible for nontrivial changes in various grammatical subsystems in numerous languages. The claim that such changes are impossible is therefore wrong; and so is the related claim, common in generative studies of language change, that very young children who are engaged in first-language acquisition are the sole initiators of significant linguistic change, because very young children lack the social maturity to participate in group decisions about language change.

It must be emphasized that in this context 'linguistic change' refers to the innovation itself, not to the spread of the innovation to later generations and/or to other adult members of the speech community. It is not necessary, therefore, to argue for the ability of an adult to change the grammar s/he internalized in early childhood; all that is necessary is that the adult produces systematically innovative linguistic structures that comprise part of the input for the next generation's language acquisition. Because they are systematic and consistent, these innovative structures must somehow form part of the adult's linguistic competence; but since an adult's competence certainly includes a variety of registers and often includes diverse dialects and even separate languages, it is unnecessary to assume that intentional innovation of new structures means a replacement of early-acquired structures by the new ones. The new structures might be additions to the adult's repertoire, not replacements of earlier structures (even if the earlier structures are no longer used at all).

It is easy to find trivial examples of intentional change. The most obvious instances are lexical innovations, for example, slang terms, which have a notoriously rapid rate of turnover. In northern California college slang of ca. 1960, a synonym for *crazy* in its slang sense was *napa*, named for a well-known California mental institution in Napa Valley. In the 1990s slang spoken by inmates in the Pennsylvania state penitentiary at Pittsburgh, a synonym for *crazy* was *snoo* – from the acronym SNU of the Special Needs Unit, where mentally disturbed inmates were housed. Similar examples can be multiplied endlessly.

Other lexical changes have at least minor structural effects, such as the replacement, by most Modern English writers, of generic *he* with *he or she, s/he*, or (though not in Standard English) *they* with a singular referent. The deliberate and highly politicized nature of this change is underlined by Roger Lass (1997: 368), who defends his use of generic *he*:

> 'I can't not use it (without a conscious decision of a type not at all characteristic of "normal" change) and still be speaking "my own language." Like all normal speakers, I am burdened by the historically given.'

This raises an important question: What is 'normal change,' and is there any practical way to exclude intentional changes from its domain, or any theoretical justification for doing so? If intentional change were impossible, it would make sense to call it nonnormal; but if it were impossible, speakers would be unable to abandon generic *he* in any case. These questions cannot be addressed directly here, but it seems extremely unlikely that every intentional change will be identifiable after the fact, and it is therefore risky to assume any sharp distinction between a normal change and a nonnormal change.

Perhaps the largest group of intentional changes is to be found in processes of language planning (see, e.g., Daoust, 1997 for a survey of language planning, or any of the over 57 000 Google hits on the subject). Language academies often focus heavily on vocabulary, but some planners recommend structural changes, especially during processes of standardization. Early in the 20th century, for instance, the Estonian language reformer Johannes Aavik introduced new structures as well as new words into Estonian, and some of these were adopted by Estonian speakers as regular features of their language; one example is the replacement of phrasal suplerlatives (as in English *more handsome*) by one-word superlatives (as in English *handsomest*) (thanks to Ilse Lehiste for this example; see Saagpakk, 1982: lxxix–lxxxvii and Oksaar, 1972 for further examples).

Another prominent motivation for deliberate linguistic changes is a concern to emphasize a group's identity vis-à-vis some other group(s). (This is sometimes also a type of language planning, though it is rarely part of an official governmental effort.) The

clearest cases result directly or indirectly from language contact.

Ma'a (Mbugu), the famous mixed language of Tanzania, is spoken in a speech community in which all of its current speakers are also fluent in at least one Bantu language, Pare. Ma'a has a voiceless lateral fricative, inherited from Cushitic; Pare, like other Bantu languages of the region, has no such phoneme. Maarten Mous notes that modern Ma'a speakers insert this lateral fricative into Pare words when speaking Pare, as a means of emphasizing their differentness from Pare and the difficulty of their Ma'a language (1994: 199) – an instance of deliberate phonological change. A comparable but more dramatic example, this time of morphological rather than phonological distortion, comes from Bougainville Island, Papua New Guinea, as described by Donald Laycock (1982). As has been reported from elsewhere in New Guinea, speakers of the Uisai dialect of Buin apparently decided that their speech was too similar to that of their neighbors; without such a decision, followed by deliberate action, it is difficult or impossible to explain the change that distinguishes their dialect: all nouns that are masculine in other dialects of Buin, with all their elaborate gender agreement marking patterns throughout the sentence, are feminine in Uisai, and vice versa. This cannot have been a gradual, incremental change, and it surely was not the result of very young children's innovation.

A sociolinguistically similar example is found in Lambayeque Quechua, a variety of Quechua spoken in Peru (and, by report, in some other varieties of Quechua as well). Speakerers of Lambayeque Quechua told a visiting linguist (Dwight Shaver, as reported to David Weber, p.c.) that they felt that they spoke too much like their neighbors, so they distorted their vocabulary by reversing the order of certain vowel-consonant sequences, replacing, e.g., *yawar* with *yawra, kabalta* with *kablata*, and *-taq* with *-tqa*.

More thoroughgoing lexical distortion has been reported from widely separated localities, occasionally with significant enough results to produce (arguably) a new language. An especially striking example was reported from Baluchistan, in what is now Pakistan, in 1913 (Bray, 1913: 139–140). Sweeping changes that included such devices as metathesis and the addition of prefixes and suffixes made this language unintelligible to nonmembers of the speech community:

> 'Mōkkī, the cant of the Lōṛīs, ... is an artificial jargon, which the Lōṛīs have mechanically invented on the basis of the language of the people among whom they live, and which they more especially employ when they want to keep their meaning to themselves ... And yet so universally and successfully is the jargon used, that it seems doubtful whether its artificiality suffices to debar it from being classed as a language ... it is at any rate acquired naturally and as a matter of course by Lōṛī children; it is no longer, it would seem, simply a secret patter; it is becoming a language for the home-circle ... both Brahui and Baloch admit freely that Mōkkī is beyond them.'

With changes in the vocabulary so massive that people who had previously understood the group's language were no longer able to do so, it is difficult to classify Mōkkī simply as a changed later form of any earlier language and thus as a member of a family tree.

There are, however, clearer examples of the disruption of language transmission, understood as the transmission of a complete language, including all grammatical subsystems, from one generation to the next, with only relatively minor incremental changes. Some of these involve intentional change in the creation of bilingual mixed languages, in contact situations in which a new (sub)ethnic group arises. Bilingual mixed languages like Michif (spoken primarily in Manitoba and North Dakota) and Mednyj Aleut (once spoken on Mednyj, or Copper, Island off the far eastern coast of Russia) cannot have arisen gradually, and they almost certainly emerged at least in part through deliberate, conscious decision. Michif arose as the language of a new mixed-blood White/Algonquian population that coalesced into a separate speech community, culturally and legally distinct from both source populations. Michif consists of Cree (Algonquian) verbal lexicon and sentence structure, including Cree phonology and fully elaborated Cree morphology, combined with French noun phrases in which the lexical items, the phonology, the morphology, and the syntax are all from Canadian French (see, e.g., Bakker and Papen, 1997). There is a bit of leakage from the Cree parts into the French noun phrases – Cree demonstratives and a Cree obviative suffix – but otherwise the two components are kept quite separate. Michif must have been created by fluent bilinguals; it cannot have been created by very young children, because children learning two first languages have not been found to create such mixtures. The language's date of origin is not known precisely, but it must have arisen in a short period of time.

Mednyj Aleut is (or was – it had almost vanished at last report) the language of a mixed-blood Aleut/Russian population that arose as a result of the Russian fur-seal trade that flourished around the Aleutian Islands during much of the 19th century. The language is mostly Aleut with some Russian admixture (loanwords and perhaps also some structural borrowings), except for the finite verb morphology,

which has been imported intact from Russian (this characterization of the language as a whole is somewhat controversial; see Thomason, 1997 and references cited there for description and discussion). Like the Michif speech community, the creators of Mednyj Aleut were a mixed-blood population, resulting from unions of Russian traders and Aleut women. The mixed-bloods held a distinct economic and social position on Mednyj Island. Here too the exact date of origin is unknown, but it must have been considerably less than a hundred years, because the conditions that gave rise to it were not present before 1826, when Aleuts were first settled on the island by Russian seal fur traders, and they were no longer present much beyond 1867, when the Russian American Company vanished and most Russians therefore left the island.

The conclusion is that adults are not only capable of inventing new words and new meanings for old words and then adding the innovative forms to their language or replacing old words with new ones; and they are not only able to modify a few fairly minor grammatical rules. They are also able to implement much more elaborate changes in their language, including massive lexical distortion and massive structural change as well. In the most extreme cases – given limits on the time available for the changes and/or on the possibilities for gradual incremental change – a significant component of these changes is almost certainly deliberate, conscious decisions by the speakers who initiate the changes; in a few cases, these decisions are actually articulated by the speakers. It is likely that the most successful intentional structural changes occur in small, compact speech communities. Exceptions may be found in some instances of national language planning, including language standardization, but nothing as elaborate as the changes that produced Uisai, Mōkkī, Michif, and Mednyj Aleut is known (at least to me) from successful language-planning efforts. Except for language planning, the literature on deliberate changes is sparse. Some discussion and references to particular examples can be found in Thomason (2001, Chaps. 6 and 8 and Thomason, 1999).

Bibliography

Bakker P & Papen R A (1997). 'Michif: a mixed language based on Cree and French.' In Thomason S G (ed.) *Contact languages: a wider perspective.* Amsterdam: John Benjamins. 295–363.

Bray D de S (1913). *Census of India, vol. 4: Baluchistan.* Calcutta: Superintendent Government Printing, India.

Daoust D (1997). 'Language planning and language reform.' In Coulmas F (ed.) *The handbook of sociolinguistics.* Oxford: Blackwell. 436–452.

Lass R (1997). *Historical linguistics and language change.* Cambridge: Cambridge University Press.

Laycock D C (1982). 'Melanesian linguistic diversity: a Melanesian choice?' In May R J & Nelson H (eds.) *Melanesia: beyond diversity.* Canberra: Australian National University Press. 33–38.

Mous M (1994). 'Ma'a or Mbugu.' In Bakker P & Mous M (eds.) *Mixed languages: 15 case studies in language intertwining.* Amsterdam: Institute for Functional Research into Language and Language Use (IFOTT). 175–200.

Oksaar E (1972). 'Bilingualism.' In Sebeok T A (ed.) *Current trends in linguistics, vol. 9: Linguistics in western Europe.* The Hague: Mouton. 476–511.

Saagpakk P F (1982). *Estonian-English dictionary.* New Haven: Yale University Press.

Thomason S G (1997). Mednyj Aleut. In Thomason S G (ed.) *Contact languages: a wider perspective.* Amsterdam: John Benjamins. 448–468.

Thomason S G (1999). 'Speakers' choices in language change.' *Studies in the Linguistic Sciences 29(2),* 19–43.

Thomason S G (2001). *Language contact: an introduction.* Edinburgh and Washington, DC: Edinburgh University Press and Georgetown University Press.

Language Change: Probabilistic Models

K Zuraw, University of California, Los Angeles, CA, USA

Properties of a Probabilistic Model of Language Change

In many areas of linguistics it makes sense to study the mental grammar of the individual in isolation. Language change, however, cannot be understood solely as something that happens within an individual (even if individuals' grammars do change over their lifetimes), but rather must be understood as something that happens in a speech community, through the accumulation of interactions between individuals over time. Each of these interactions has a probabilistic character: which individuals will interact, which variant the speaker will produce, and how the listener's mental grammar and lexicon will be affected.

Because language change is not instantaneous at the population level, any well developed model of

which has been imported intact from Russian (this characterization of the language as a whole is somewhat controversial; see Thomason, 1997 and references cited there for description and discussion). Like the Michif speech community, the creators of Mednyj Aleut were a mixed-blood population, resulting from unions of Russian traders and Aleut women. The mixed-bloods held a distinct economic and social position on Mednyj Island. Here too the exact date of origin is unknown, but it must have been considerably less than a hundred years, because the conditions that gave rise to it were not present before 1826, when Aleuts were first settled on the island by Russian seal fur traders, and they were no longer present much beyond 1867, when the Russian American Company vanished and most Russians therefore left the island.

The conclusion is that adults are not only capable of inventing new words and new meanings for old words and then adding the innovative forms to their language or replacing old words with new ones; and they are not only able to modify a few fairly minor grammatical rules. They are also able to implement much more elaborate changes in their language, including massive lexical distortion and massive structural change as well. In the most extreme cases – given limits on the time available for the changes and/or on the possibilities for gradual incremental change – a significant component of these changes is almost certainly deliberate, conscious decisions by the speakers who initiate the changes; in a few cases, these decisions are actually articulated by the speakers. It is likely that the most successful intentional structural changes occur in small, compact speech communities. Exceptions may be found in some instances of national language planning, including language standardization, but nothing as elaborate as the changes that produced Uisai, Mōkkī, Michif, and Mednyj Aleut is known (at least to me) from successful language-planning efforts. Except for language planning, the literature on deliberate changes is sparse. Some discussion and references to particular examples can be found in Thomason (2001, Chaps. 6 and 8 and Thomason, 1999).

Bibliography

Bakker P & Papen R A (1997). 'Michif: a mixed language based on Cree and French.' In Thomason S G (ed.) *Contact languages: a wider perspective.* Amsterdam: John Benjamins. 295–363.

Bray D de S (1913). *Census of India, vol. 4: Baluchistan.* Calcutta: Superintendent Government Printing, India.

Daoust D (1997). 'Language planning and language reform.' In Coulmas F (ed.) *The handbook of sociolinguistics.* Oxford: Blackwell. 436–452.

Lass R (1997). *Historical linguistics and language change.* Cambridge: Cambridge University Press.

Laycock D C (1982). 'Melanesian linguistic diversity: a Melanesian choice?' In May R J & Nelson H (eds.) *Melanesia: beyond diversity.* Canberra: Australian National University Press. 33–38.

Mous M (1994). 'Ma'a or Mbugu.' In Bakker P & Mous M (eds.) *Mixed languages: 15 case studies in language intertwining.* Amsterdam: Institute for Functional Research into Language and Language Use (IFOTT). 175–200.

Oksaar E (1972). 'Bilingualism.' In Sebeok T A (ed.) *Current trends in linguistics, vol. 9: Linguistics in western Europe.* The Hague: Mouton. 476–511.

Saagpakk P F (1982). *Estonian-English dictionary.* New Haven: Yale University Press.

Thomason S G (1997). Mednyj Aleut. In Thomason S G (ed.) *Contact languages: a wider perspective.* Amsterdam: John Benjamins. 448–468.

Thomason S G (1999). 'Speakers' choices in language change.' *Studies in the Linguistic Sciences 29(2),* 19–43.

Thomason S G (2001). *Language contact: an introduction.* Edinburgh and Washington, DC: Edinburgh University Press and Georgetown University Press.

Language Change: Probabilistic Models

K Zuraw, University of California, Los Angeles, CA, USA

Properties of a Probabilistic Model of Language Change

In many areas of linguistics it makes sense to study the mental grammar of the individual in isolation. Language change, however, cannot be understood solely as something that happens within an individual (even if individuals' grammars do change over their lifetimes), but rather must be understood as something that happens in a speech community, through the accumulation of interactions between individuals over time. Each of these interactions has a probabilistic character: which individuals will interact, which variant the speaker will produce, and how the listener's mental grammar and lexicon will be affected.

Because language change is not instantaneous at the population level, any well developed model of

language change must be probabilistic at least in the sense of specifying how an individual behaves when surrounded by variation and how that behavior in turn affects the rest of the population. In the simplest model, a learner could simply adopt the grammar of the first other speaker encountered (assuming that the encounter provides enough information for the learner to identify that grammar). In that case, the model is probabilistic only in determining what type of speaker the learner encounters. But more realistic models require learning algorithms that can handle the learner's variable environment.

It can be difficult (or misleading) to try to understand how such a model will behave without being mathematically explicit about how the model works, and examining its behavior under varying conditions, whether analytically – by studying properties of equations – or through computer simulation when direct analysis is infeasible. Being explicit requires simplifying reality somewhat, and depending on researchers' interests, models are typically more realistic in one area than in others.

A model of language must specify what the internal state of an individual looks like. First, does an individual have a single grammar that produces no variation, a single grammar that produces probabilistic variation, or multiple grammars each associated with a probability of use? ('Grammar' here is used loosely to include not only a language's phonology, morphology, and syntax, but also possibly its lexicon.) Second, how is a grammar represented? Possibilities that have been proposed include a vector of parameter settings, a connectionist network, a cloud of exemplars, or a probabilistic ranking of Optimality-Theoretic constraints.

It is also vital to specify how an individual's grammar is arrived at during initial acquisition, and how it may change during adulthood; the choice of learning or update algorithm of course depends on the nature of the grammar. Once a learning algorithm is specified, it is still necessary to consider whether there is a critical period after which learning ceases, and if so how much data the learner is exposed to during that period; if there is no critical period, the model must specify how plastic the grammar remains as the individual ages.

Moving to the level of the population, the model must specify who interacts with whom. Do children learn only (or preferentially) from their parents? Do all individuals interact equally with all members of the population, or is the population structured, whether geographically or socially, so that some interactions are more likely than others? Does the listener in an interaction treat all data the same way, or is an utterance's influence a function of the speaker's (and possibly the listener's) age, the strength of the social connection between the two individuals, or even the social status of the speaker? Does the speaker's choice of utterance similarly depend on characteristics of the listener? Properties of the population itself must also be decided, such as size, nature of geographic and social structure (if any), birth and death rates, and whether these differ as functions of grammar, geographic location, or status in the social network (see Milroy, 1987 on the role of social networks in variation and change).

Probabilistic Grammars

It has been well documented that when a linguistic change is in progress, productions vary, at rates that increasingly favor the innovative variant, not just across individuals but also within individuals (see Labov, 1994, 2001; Kroch, 2001 and references therein). Weinreich *et al.* (1968) proposed to model this with **variable rules**, phonological rules that apply not obligatorily, but with a certain probability, which can change from generation to generation (see also Paolillo, 2001). Little work has been done on modeling acquisition in this framework.

Probabilistic Parameter Settings

In syntax, Kroch (1989 and subsequent work by Kroch and colleagues; see Kroch, 2001 for references) has attached probabilities to parameter settings. While change is in progress, individuals have nonzero probabilities associated with both settings of some parameters.

Learning algorithms have been proposed for parameter setting in variable environments, but these algorithms generally do not result in a variable grammar. Niyogi and Berwick (1995), for example, consider the behaviors of variants of Gibson and Wexler's (1994) Triggering Learning Algorithm when the learning environment contains speakers with different grammars; but even in the face of contradictory data, the learner chooses a single grammar. In Briscoe's approach (2000 and references therein), the learner also arrives at a single vector of parameter settings, but during learning maintains probabilities for those settings, based on both their prior probabilities (universally given) and their probabilities in light of the learning data; these probabilities determine how easily a parameter setting can be changed in response to further learning data.

Competing Grammars

Rather than having variation within a single grammar, it is also possible to imagine that speakers can maintain multiple, independent grammars, each

having some probability of use. (As a model of diglossic or bilingual competence, such multiple grammars would need to be associated with additional information about their appropriateness in different social settings.) Yang (2000) models acquisition of competing grammars in a reward-penalty scheme. When the learner encounters an utterance, a grammar is selected at random according to its current probability. If the grammar can parse the utterance, its probability is increased; otherwise, its probability is decreased.

Competing full grammars permit a wider range of variation types than grammars with probabilistic rules, parameter settings, or constraint rankings. For example, in a grammar with parameters A and B, each having possible settings 0 and 1, a speaker who maintains separate grammars could use the grammars (0,0) (A = 0 and B = 0) and (1,1), each at probability 0.5. But if the speaker associates probabilities not with grammars but with individual parameter settings, this will be impossible: if (0,0) and (1,1) are both used, then the parameter settings A = 0, A = 1, B = 0, and B = 1 all must have nonzero probabilities, so the speaker will sometimes use (0,1) and (1,0) as well. The situation is similar for variable rules and constraint rankings. It should be noted that the term 'competing grammars' is often applied to cases of a single grammar with variable parameter settings.

Probabilistic Constraint Ranking

Probabilistic ranking of Optimality-Theoretic (OT) constraints has also been applied to both phonological and syntactic change: in this version of OT, a grammar consists not of a linear ranking of constraints, but of a probability distribution over possible constraint rankings. One such model is Boersma's (1998) stochastic OT, in which each constraint is associated with a ranking value; at the time of generating an utterance, random noise is added to each ranking value, and the resulting perturbed values are used to rank the constraints linearly. In this theory, learning is accomplished by Boersma's Gradual Learning Algorithm.

Adaptive Rules

An alternative to probabilistic grammars during language change is the theory of adaptive rules (Andersen, 1973), in which each individual has a non-variable grammar. But, if that grammar represents an innovative departure from the current standard, the speaker develops a layer of rules to make her utterances conform to community norms. Such a speaker's adult utterances are just like a conservative speaker's; the difference is that the innovative speaker is 'tolerant' of the same innovation in children. These children, facing less opposition, would then develop their adaptive rules later in life – meanwhile producing utterances with the innovation, which might serve as learning data to younger learners – and in turn be even more tolerant of innovation. This process continues until the adaptive rules are used only in a special, conservative style associated with the old, and eventually the adaptive rule is eliminated entirely. This theory rests on the assumption that a child learns not just from adults' utterances, but also from adults' reactions to the child's own speech; most models do not share this assumption.

Probabilistic Categories

The probabilistic models of grammar described above rely on discrete categories, such as nouns vs. verbs, or [+high] vs. [−high] vowels. In some models of language change, however, category membership is itself probabilistic.

Exemplar-Based Models

Pierrehumbert (2002) has explored certain types of phonological change within an exemplar-based framework. In exemplar theory, tokens of encountered speech are mapped onto a similarity space with dimensions such as duration and formant values, and possibly time. Not every token is necessarily stored separately, because the similarity space is granular, or discretized into 'bins' called exemplars. When a new token is encountered, the strength of the exemplar it belongs to is augmented (this could be done either incrementally or by giving an exemplar maximum strength whenever a token is mapped to it). Countervailingly, exemplar strength decays over time. The learner forms categories – corresponding, for example, to members of the ambient language's vowel inventory – by identifying clusters of exemplars. Each category, then, is made up of many exemplars of varying strengths. In perception, a category label is attributed to an incoming stimulus according to the number of exemplars from each category that are similar to it, with a weighting in favor of stronger exemplars. In order to produce an instance of a category, an exemplar is chosen at random, but again with weighting in favor of stronger exemplars, and neighbors of the chosen exemplar contribute to the resulting production in proportion to their strengths; random noise may also be added. Lexical entries, moreover, contain weights that are applied to exemplars in production, so that one word may 'prefer' to be realized with a vowel of long duration, while another prefers short duration.

Pierrehumbert models Joan Bybee's discovery that lenitory phonological changes (those involving

gestural reduction) progress more quickly in high-frequency words. For example, English optional schwa reduction/deletion is more advanced in high-frequency words such as *every*, *camera*, *memory*, *family* than in low-frequency words such as *mammary*, *artillery*, *homily* (see Bybee, 2001 and references therein). Bybee proposes that the reason for the phenomenon is that lenition is imposed every time a word is used and the lexical entry (of the listener and/or speaker) is updated in response. Since this happens more often in high-frequency words, they change more rapidly. Crucial to this explanation is allowing lexical entries to have a high degree of phonetic concreteness: in the English schwa example, the lexical entry cannot simply be a string of phonemes, either containing or not containing a medial vowel. Rather, it must somehow contain information about the typical duration range (including zero) of the potential schwa vowel, such as through lexical weighting of schwa exemplars. Pierrehumbert's simulations, using repeated rounds of production and perception in the exemplar model with an externally imposed constant tendency representing lenition, show that change does indeed progress more rapidly in frequent words.

Pierrehumbert proposes that a phonetic change can eventually change lexical entries at the phonological level. For example, it is plausible that words like *family*, whose phonology allows schwa deletion (sonorants surround the medial vowel) are currently nonetheless represented lexically with a medial schwa, just as are words like *attitude*. But if English schwa reduction continues to progress, words like *family* will come to weight short exemplars of schwa so heavily that their schwa productions cluster around a very short (or even zero) duration, whereas the schwa productions of words like *attitude* will cluster around a longer duration. Since, in Pierrehumbert's model, learners infer phoneme-like categories from clustering, in subsequent generations the lexical entries for the two types of words will refer to different category strings.

Connectionist Models

Connectionist models of competence have also been applied to language change, by such authors as Tabor (1994) and Johansson (1997).

Tabor's model of the grammar/lexicon is associative, with no explicit parameter settings or grammatical categories. The model is a connectionist network with three layers of nodes: input, output, and hidden. Input nodes represent lexical items, and output nodes represent word behaviors, such as taking a certain type of complement. Activation of an input node flows to strongly connected hidden-layer nodes, which in turn activate strongly connected output-layer nodes; there is no direct connection between input and output nodes. Fuzzy grammatical categories, which can be conceptualized as patterns of input-node-to-hidden-node connection weight, emerge from similar distribution. Two words that would both be classified in a traditional grammar as nouns, for example, have relatively similar distributions and therefore similarly weighted connections to the hidden units. The use of this network to model language change is discussed in Constant Rate Effects and Learning Biases below.

Constant Rate Effects

When a change, such as from SOV (subject-object-verb) to SVO word order, occurs in multiple syntactic environments, such as main and subordinate clauses, we can imagine different patterns of change. If for some reason the change originates in main clauses, SVO could replace SOV fairly quickly there, and more slowly in subordinate clauses, perhaps because the change is only gradually being generalized to environments that are less similar to the original environment of the change.

An alternative, advocated by Anthony Kroch and colleagues, is that when the syntactic theory dictates that variation in multiple environments be controlled by the same parameter (such as head-complement order), historical replacement of one variant by the other should proceed at the same rate in all environments, though it may, for poorly understood reasons, begin earlier in some environments than others. This is the 'constant rate hypothesis,' and it has been argued to be borne out by numerous case studies of syntactic change (see Kroch, 2001 and references therein).

In order to test the constant rate hypothesis, 'rate of change' must be defined. Kroch uses the logistic function $P = 1/(1 + e^{-k-st})$ to model change, where P is the probability of use of the new variant, t is time, and k and s are constant parameters that determine when the change begins and how quickly it progresses (see Contacts in the Population below for derivation of this function). Algebra transforms this equation into $[\ln P/(1 - P)] = k + st$; we see that the logistic transform, or logit $[\ln P/(1 - P)]$ is equal to $k + st$, a linear function of t with slope (steepness) s and intercept (starting point) k. Kroch takes the slope s to be the rate of change; thus s is expected to be constant across environments.

Tabor (1994), however (see Connectionist Models, above), views constant-rate change as merely a special case of *frequency linkage*. When a word or construction's distribution changes, the input-output

relationship cannot be directly changed in response, since there are no input-to-output connections; only the input-to-hidden and hidden-to-output connection weights can change. When the hidden-to-output weights change because one word or construction shifts in distribution, any word or construction with a similar pattern of input-to-hidden connection weights is necessarily 'dragged along' in the change to some extent. That is, words and constructions undergo similar changes in distribution to the degree that their previous distributions were similar. Constant-rate cases, for Tabor, are simply extreme cases in which two constructions' initial distributions are highly similar.

Learning Biases

Why are speakers biased to adopt some linguistic innovations, so that they spread, rather than faithfully imitating whatever they encounter equally, so that new variants fail to spread? Attempts to model such bias have mostly taken one of two approaches. The first approach simply assumes that the new variant is preferred by speakers, perhaps because it carries social prestige of some kind. This approach is inspired by the sociolinguistic research into changes in progress (most phonological) of William Labov and colleagues (see Labov, 1994, 2001 and references therein; a truly Labovian model would require more than merely tagging one variant as preferred).

The second approach, which has been more extensively developed, explores cases in which one variant might have a structural advantage over another, usually because it generates fewer ambiguous utterances. (See Clark, 2004 for another approach – a processing filter that skews the data actually usable for learning.) A commonly used example is word order. Niyogi and Berwick (1995), for example, model the change from Old French verb-second (V2) word order to Modern French subject-verb-object (SVO) order. In their simulation, a grammar is represented as a vector of (non-probabilistic) binary parameter settings, and language acquisition consists of determining those settings based on input utterances. Acquisition would be easy if every utterance uniquely identified the grammar it came from, but there are many utterances that could be produced by more than one grammar. For example (to use English lexical items), the utterance 'dog bites man' is consistent with either a V2 or an SVO grammar, whereas 'suddenly bites dog man' is not consistent with SVO and 'dog suddenly bites man' is not consistent with V2. We can imagine various strategies that a learner might use in response to an ambiguous utterance such as 'dog bites man': ignore the utterance, increment the probability of every grammar consistent with the utterance, or randomly choose one grammar consistent with the utterance and augment its probability. A learner need not even recognize the ambiguity: under a Triggering-Learning approach, the learner changes grammars only if the current grammar fails to parse an incoming utterance; an ambiguous sentence will cause no change if the learner's current grammar is consistent with it.

Under all these strategies, a grammar is at a disadvantage if it produces a higher proportion of ambiguous sentences than a competing grammar does. For example, consider a population in which 20% use innovative L_1 and 80% use conservative L_2. If 80% of L_1's utterances are unambiguous, compared with 10% of L_2's, then for every 100 utterances, 16 are triggers for L_1 and only 8 are triggers for L_2 (76 are ambiguous). Although L_1 is rarer in the population, from the learner's point of view there are more utterances coming from L_1 than from L_2. L_1 will therefore inexorably encroach on L_2.

Working in the framework of Stochastic Optimality Theory, Jäger (2003) has also modeled learning bias as a source of language change. In Jäger's variant of Optimality Theory, learning data are form-meaning pairs. The learner checks whether its current grammar (probabilistically) chooses the observed meaning as optimal given the observed form, but also whether its current grammar chooses the observed form given the observed meaning. If either type of mismatch occurs, the learner adjusts constraint ranking values accordingly, as in Boersma's (1998) Gradual Learning Algorithm. Jäger simulates the behavior of this system through repeated rounds of production and learning, and finds that when one type of form is penalized by the constraint inventory more than another (regardless of ranking), the 'better' form becomes associated to the more frequent meaning. This is because if the system begins unbiased, it will tend to use the 'better' form for both meanings. If the learner's grammar picks the wrong meaning as optimal, it must demote any constraint that disfavors the observed form-meaning pair. The more frequent meaning is of course encountered more often, so the constraint against associating that meaning to the better form gets demoted more vigorously, and the frequent meaning increasingly is produced with the better form. Jäger applies this model to case-marking systems dependent on animacy; corpus data on real languages reveals pragmatically motivated asymmetries, such as that sentences in which an animate actor acts on an inanimate patient are common. Jäger is able to produce the result that zero case-marking (which he takes to be favored) tends to become associated with

the common combinations of animacy and thematic role.

Tabor (1994) approaches biases from a different angle, focusing on possible trajectories from one grammar to another. Tabor asks why English *be going to*, on its journey from meaning solely 'be traveling towards' in Old English to acting as a future auxiliary today, passed through the stages it did. *Be going to* took on new traits one after the other, rather than simultaneously increasing the frequency of all those traits: first it began to appear with verb-phrase complements (rather than exclusively location complements) but with an intentional meaning and with motion plausibly involved (as in modern *we're going to fight*), then with no motion (*he's going to say something*), then with lack of intention (*he's going to be fired*), then with non-sentient subjects (*the chair is going to break*), and finally with dummy subjects (*it's going to rain*). In Tabor's connectionist model (see Exemplar-Based Models, above), categories are represented in a continuous space, so if an item such as *be going to* moves from one category to another (by changing its input-to-hidden weights), it may pass through other categories on the way. The connection weights in the network determine what trajectories through this space are possible, and Tabor shows how a plausible model of verb categories would cause *be going to* to pass through the category of equi verbs (which take verb-phrase complements but require sentient subjects) on its way from motion verb to raising verb.

Probabilistic Aspects of Reanalysis

Reanalysis has been argued to result from frequency shifts that cause a structure to become associated with a context that admits or encourages the new analysis (e.g., Lightfoot, 1991). Others have argued, based on case studies, that reanalysis precedes the structural change (Frisch, 1994, Pintzuk, 1995); in those cases, it is unclear what triggers the reanalysis in the first place.

Fontana (1993) explores in detail a case in which one syntactic change affects a construction's environment and thus causes another syntactic change. Object pronouns in Old Spanish appeared in second position (enclitic on the first word of the clause), but in Modern Castilian Spanish they are verbal proclitics. Loss of topicalization in Old Spanish led to an increase in clauses beginning with the object pronoun; assuming that object pronouns were not able to stand on their own phonologically, these sentence-initial pronouns would have been procliticized to the following word. Another change, loss of I-to-C movement, meant that the following word was often a verb. As the pronouns appeared as proclitics on verbs more and more often, they were reanalyzed to be permitted in that position only, whether sentence-initial or not.

Although the details of the analysis depend on issues in syntactic theory, if the basic outline is right, the Spanish situation can be seen as one in which an increasing number of forms become ambiguous, being consistent with either the existing grammar or an innovative one. The learning algorithm should then predict under what conditions, if any, the innovative grammar can take over (see Learning Biases, above).

Deriving S-Shaped Change

It has long been observed that many historical developments follow an S-shaped pattern of change (see Denison, 2002 for discussion). At first, a new variant is rarely used, and its use increases slowly; in the middle period of the change, the new variant gains ground more and more quickly until the two variants' frequencies are about equal, after which point the change slows down; in the late period of the change, the older variant disappears more and more slowly and may even persist, at very low frequencies, indefinitely. An S-curve is illustrated in Figure 1.

Contacts in the Population

Many explanations involving probabilistic interactions among individuals in the speech community have been proposed for this pattern of change. One possibility, mentioned by Labov (1994: 66), is that the S shape of change simply reflects the time-course of social contacts. If we assume, for simplicity's sake, that each individual converts to the new variant on first encountering it, a change that originates in one member of a population will still take time to spread. At the first time-step, only those who interact directly with the innovator 'catch' the new variant, and its spread is slow. But as the number of speakers with the innovative variant increases, so does the number of speakers in contact with them at every time-step, so the spread accelerates, until more than half of the speech community uses the new variant, at which point the spread slows because there are fewer and fewer speakers left who do not already use the new variant.

To see this mathematically, suppose that there are n speakers in the population. Let x represent the number of speakers with the new variant at any time; $n - x$ is the number of speakers with the old variant. The number of possible pair-wise interactions between speakers that would result in conversion of a speaker from the old variant to the new is simply the number

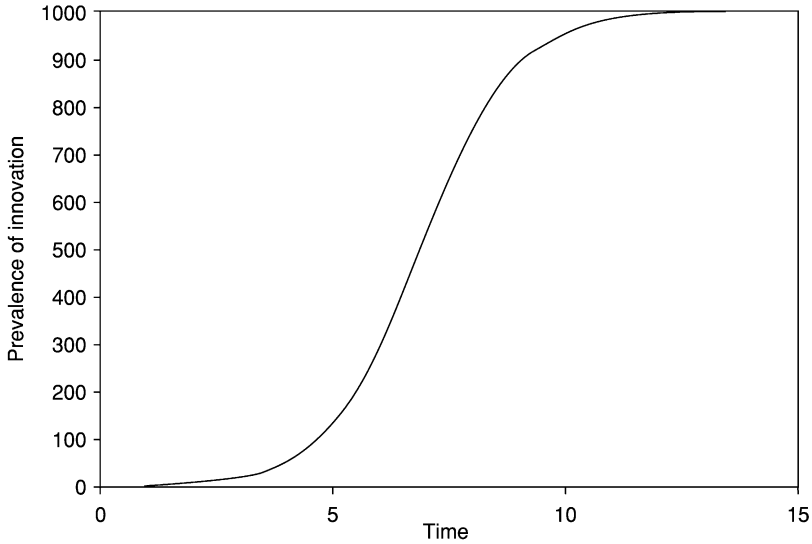


Figure 1 An S-shaped curve.

of pairs composed of one new-variant speaker and one old-variant speaker, or $x(n-x)$. But, at a given time-step, not all those interactions will take place. Let p, a number between 0 and 1, represent, for each pair of speakers, the probability that they meet at any one time-step. At each time-step, then, the increase in speakers with the new variant is $px(n-x)$. When x is small this quantity is small – change is slow. For example, with $n = 100$ and $p = 0.01$, when $x = 1$, the number of speakers converted to the new grammar is $p(n-1) = 0.99$, just under one speaker. When x reaches $n/2 = 50$, the rate of change reaches its maximum, $pn^2/4 = 25$. When the change is almost complete, $x = 99$, the rate of change is again $p(n-1) = 0.99$.

If we examine the proportion u ($= x/n$) of the population that has the innovation, the rate of change at every timestep can be written $au(1-u)$, where a ($=pn^2$) plays a role similar to p's in determining how frequent contact is. To obtain u as a function of time we must solve the differential equation $du/dt = au(1-u)$. The solution is of the form $u = 1/(1 + e^{-C-at})$ where C is a constant. This is the logistic function seen in Constant Rate Effects, above. Setting C so that the population starts with 1% of members having the new variant, we obtain the plot for $a = 0.5$ shown in Figure 2.

An S-shaped pattern of spread is still derived if speakers' adoption of the new variant is probabilistic (e.g., each contact with the new variant only increments a speaker's use of it), or if the speech community is organized into a social network, with contacts between speakers being governed by social bonds instead of evenly distributed.

As discussed above in Learning Biases, many authors have constructed models in which the tendency of one linguistic variant to spread emerges from learning, rather than simply being imposed, and there have been simulation studies (see Niyogi, 2004 and references therein) investigating under what learning conditions the new variant spreads according to an S shape.

A Caution About Rates of Change

Some authors have cautioned that, because of stylistic conservatism, change may appear, in the written record, to be slower than it really was. Shi (1989) argues that although the rise of the aspectual particle *le* in Mandarin Chinese appears, from the written record, to take about 1000 years, it was actually much more abrupt. Shi argues is that even after the rise of *le* was completed in the spoken language, the written language contained a mix of contemporary and classical styles even within documents, so that the rate of *le* use appears artificially low. To obtain the vernacular rate of *le* use from the written record requires factoring out the rate of classical-style use. To do this, Shi tracks the frequency of the classical copula/interjective *ye*. In true classical texts, there are 8 occurrences of *ye* per 1000 characters, so Shi infers that if a text has $n \leq 8$ occurrences of *ye* per 1000 characters, the probability that a character in that text should be attributed to the classical style is n/8. When the number of *le*s is plotted per 1000 putatively non-classical characters instead of per 1000 raw characters, the change is completed in at most 200 years. When change is so abrupt, it becomes difficult to determine whether the S shape is a good fit.

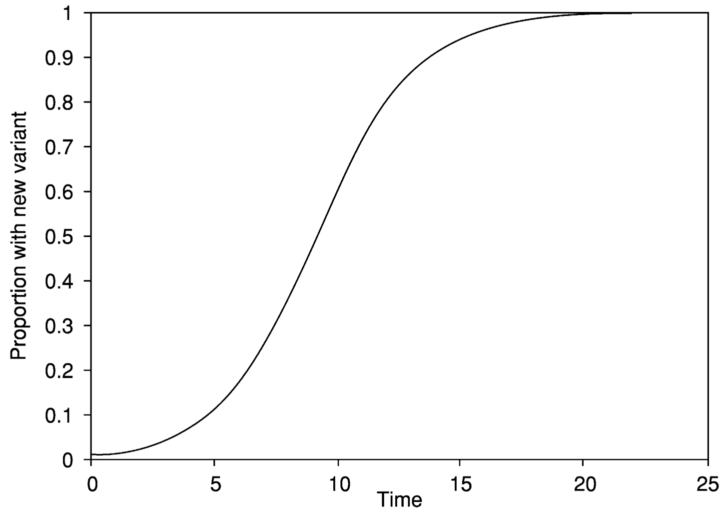


Figure 2 Logistic function.

Model Dynamics

Niyogi (forthcoming) demonstrates some unexpected properties of population-level models of language change that would be difficult to intuit. Let us consider one example that uses a cue-based learning algorithm. Cue-based learning (Dresher and Kaye, 1990) relies on the idea that certain sentences generated by a grammar could not come from any other grammar, and thus serve as unambiguous *cues* to the learner as to which grammar to choose. (On a smaller scale, a cue may also be partial, demonstrating unambiguously the setting of some parameter but being irrelevant to other aspects of the grammar.) For simplicity, assume that the learner faces a binary choice between two grammars, L_1 and L_2. Cue-based learning typically establishes one choice as the default (say L_2) by specifying that unless some minimum rate of L_1 cues is encountered, the learner will choose L_2. Varying the probability p that a speaker of L_1 produces a cue, we can see how the proportion of the population that eventually acquires L_1 (given infinite time) is affected. Niyogi finds that for small values of p, the population moves towards total L_2 use. In dynamical-systems terms, 0% L_1 use is the only stable fixed point for these low values of p – all other states will eventually move to 0% L_1 and stay there. If we examine increasing values of p, at some critical point (determined by the number of utterances in the critical period and the rate of cues that the learner requires before choosing L_1) there emerges a second stable fixed point of 100% L_1 use, as well as an unstable fixed point somewhere between 0% and 100%. That is, at higher p, a population that has little L_1 use will still move to 0% L_1 (since there are few utterances produced from L_1), but a population with more than some minimum number of L_1 speakers will now move towards 100% L_1 use and stay there; additionally, there is some intermediate proportion of L_1 use that can persist, but if the population diverges slightly from that value, it will be pulled towards 0% or 100% L_1 use. This is an example of a 'bifurcation': a small change in a parameter (here, p) results in a sudden change in the behavior of the system. Niyogi suggests that the potential for bifurcations helps address the actuation problem of Weinreich *et al.* (1968), i.e., why does a language change start?: a small change in the system, perhaps brought about by external forces such as population movements or cultural changes that change the frequency with which certain types of propositions are expressed, could lead to dramatic linguistic change. The calculations that Niyogi performs to arrive at the bifurcation result demonstrate the importance of making the model fully explicit – otherwise, its behavior cannot be seriously studied.

See also: Corpus Linguistics; Formal Models and Language Acquisition; Human Language Processing: Connectionist Models; Labov, William (b. 1927); Principles and Parameters Framework of Generative Grammar; Syntax: Optimality Theory; Weinreich, Uriel (1926–1967).

Bibliography

Andersen H (1973). 'Abductive and deductive change.' *Language 49,* 765–793.

Boersma P (1998). *Functional phonology: formalizing the interactions between articulatory and perceptual drives.* The Hague: Holland Academic Graphics.

Briscoe E J (2000). 'Evolutionary perspectives on diachronic syntax.' In Pintzuk S, Tsoulas G & Warner A (eds.)

Diachronic syntax: models and mechanisms. Oxford: Oxford University Press. 75–108.
Bybee J (2001). *Phonology and language use.* Cambridge, UK: Cambridge University Press.
Clark B (2004). *A Stochastic Optimality Theory approach to syntactic change.* Ph.D. diss., Stanford University.
Denison D (2002). 'Log(ist)ic and simplistic S-curves.' In Hickey R (ed.) *Motives for language change.* Cambridge, UK: Cambridge University Press.
Dresher E & Kaye J (1990). 'A computational learning model for metrical phonology.' *Cognition 34,* 137–195.
Fontana J M (1993). *Phrase structure and the syntax of clitics in the history of Spanish.* Ph.D. diss., University of Pennsylvania.
Frisch S (1994). 'Reanalysis precedes syntactic change: evidence from Middle English.' *Studies in the Linguistic Sciences 24,* 187–201.
Gibson E & Wexler K (1994). 'Triggers.' *Linguistic Inquiry 25,* 407–454.
Hinskens F, van Hout R & Wetzels W L (eds.) (1997). *Variation, change and phonological theory.* Amsterdam: Benjamins.
Jäger G (2003). 'Simulating language change with functional OT.' In Kirby S (ed.) *Proceedings of language evolution and computation workshop/course at ESSLLI.* Vienna: European Summer School in Logic, Language and Information. 52–61.
Johansson C (1997). *A view from language: growth of language in individuals and populations.* Lund, Sweden: Lund University Press.
Kroch A (1989). 'Reflexes of grammar in patterns of language change.' *Language Variation and Change 1,* 199–244.
Kroch A (2001). 'Syntactic change.' In Baltin M & Collins C (eds.) *The handbook of contemporary syntactic theory.* Malden, MA: Blackwell. 699–729.
Labov W (1994). *Principles of linguistic change 1: Internal factors.* Cambridge, MA: Blackwell.
Labov W (2001). *Principles of linguistic change 2: Social factors.* Malden, MA: Blackwell.
Lightfoot D (1991). *How to set parameters: evidence from language change.* Cambridge, MA: MIT Press.
Milroy L (1987). *Language and social networks* (2nd edn.). Oxford: Blackwell.
Niyogi P (forthcoming). *The computational nature of language learning and evolution.* Cambridge, MA: MIT Press.
Niyogi P & Berwick R (1995). 'The logical problem of language change.' MIT Artificial Intelligence Laboratory Memo No. 1516.
Paolillo J C (2001). *Analyzing linguistic variation: statistical models and methods.* Chicago: University of Chicago Press.
Pierrehumbert J (2002). 'Word-specific phonetics.' In Gussenhoven C & Warner N (eds.) *Laboratory phonology VII.* Berlin: Mouton de Gruyter. 101–140.
Pintzuk S (1995). 'Variation and change in Old English clause structure.' *Language Variation and Change 7,* 229–260.
Shi Z (1989). 'The grammaticalization of the particle *le* in Mandarin Chinese.' *Language Variation and Change 1,* 99–114.
Tabor W (1994). *Syntactic innovation: a connectionist model.* Ph.D. diss., Stanford University.
Weinreich U, William L & Marvin H (1968). 'Empirical foundations for a theory of language change.' In Lehmann W & Malkiel Y (eds.) *Directions for historical linguistics.* Austin: University of Texas Press. 97–195.
Yang C (2000). 'Internal and external forces in language change.' *Language Variation and Change 12,* 231–250.

Language Development in Blind Children

M Pérez-Pereira, University of Santiago de Compostela, Santiago de Compostela, Spain

Introduction

The study of how blind children learn to talk is interesting for two reasons. First, the more we understand the process of language acquisition in blind children, the better we can help them to progress in their development. Second, the study of blind children can help us understand the role of vision in language development and, in turn, test theories of language acquisition. To examine the effect of vision on language development, this article concentrates on children born blind (congenital blindness); it excludes children who lost vision later in development and therefore had visual experience before suffering blindness, as well as children who have residual functional vision (visually impaired children).

Given the great individual differences that exist among congenitally blind children, it is difficult to depict a typical or characteristic portrait of these children, and generalizations must be considered with caution.

Blindness affects the development of congenitally blind children in different ways, and language development is one area minimally affected by the lack of vision.

The use and knowledge of language in blind children has been a controversial topic. The first studies on this population, carried out in the 1950s, 1960s,

Diachronic syntax: models and mechanisms. Oxford: Oxford University Press. 75–108.
Bybee J (2001). *Phonology and language use.* Cambridge, UK: Cambridge University Press.
Clark B (2004). *A Stochastic Optimality Theory approach to syntactic change.* Ph.D. diss., Stanford University.
Denison D (2002). 'Log(ist)ic and simplistic S-curves.' In Hickey R (ed.) *Motives for language change.* Cambridge, UK: Cambridge University Press.
Dresher E & Kaye J (1990). 'A computational learning model for metrical phonology.' *Cognition 34,* 137–195.
Fontana J M (1993). *Phrase structure and the syntax of clitics in the history of Spanish.* Ph.D. diss., University of Pennsylvania.
Frisch S (1994). 'Reanalysis precedes syntactic change: evidence from Middle English.' *Studies in the Linguistic Sciences 24,* 187–201.
Gibson E & Wexler K (1994). 'Triggers.' *Linguistic Inquiry 25,* 407–454.
Hinskens F, van Hout R & Wetzels W L (eds.) (1997). *Variation, change and phonological theory.* Amsterdam: Benjamins.
Jäger G (2003). 'Simulating language change with functional OT.' In Kirby S (ed.) *Proceedings of language evolution and computation workshop/course at ESSLLI.* Vienna: European Summer School in Logic, Language and Information. 52–61.
Johansson C (1997). *A view from language: growth of language in individuals and populations.* Lund, Sweden: Lund University Press.
Kroch A (1989). 'Reflexes of grammar in patterns of language change.' *Language Variation and Change 1,* 199–244.
Kroch A (2001). 'Syntactic change.' In Baltin M & Collins C (eds.) *The handbook of contemporary syntactic theory.* Malden, MA: Blackwell. 699–729.
Labov W (1994). *Principles of linguistic change 1: Internal factors.* Cambridge, MA: Blackwell.
Labov W (2001). *Principles of linguistic change 2: Social factors.* Malden, MA: Blackwell.
Lightfoot D (1991). *How to set parameters: evidence from language change.* Cambridge, MA: MIT Press.
Milroy L (1987). *Language and social networks* (2nd edn.). Oxford: Blackwell.
Niyogi P (forthcoming). *The computational nature of language learning and evolution.* Cambridge, MA: MIT Press.
Niyogi P & Berwick R (1995). 'The logical problem of language change.' MIT Artificial Intelligence Laboratory Memo No. 1516.
Paolillo J C (2001). *Analyzing linguistic variation: statistical models and methods.* Chicago: University of Chicago Press.
Pierrehumbert J (2002). 'Word-specific phonetics.' In Gussenhoven C & Warner N (eds.) *Laboratory phonology VII.* Berlin: Mouton de Gruyter. 101–140.
Pintzuk S (1995). 'Variation and change in Old English clause structure.' *Language Variation and Change 7,* 229–260.
Shi Z (1989). 'The grammaticalization of the particle *le* in Mandarin Chinese.' *Language Variation and Change 1,* 99–114.
Tabor W (1994). *Syntactic innovation: a connectionist model.* Ph.D. diss., Stanford University.
Weinreich U, William L & Marvin H (1968). 'Empirical foundations for a theory of language change.' In Lehmann W & Malkiel Y (eds.) *Directions for historical linguistics.* Austin: University of Texas Press. 97–195.
Yang C (2000). 'Internal and external forces in language change.' *Language Variation and Change 12,* 231–250.

Language Development in Blind Children

M Pérez-Pereira, University of Santiago de Compostela, Santiago de Compostela, Spain

Introduction

The study of how blind children learn to talk is interesting for two reasons. First, the more we understand the process of language acquisition in blind children, the better we can help them to progress in their development. Second, the study of blind children can help us understand the role of vision in language development and, in turn, test theories of language acquisition. To examine the effect of vision on language development, this article concentrates on children born blind (congenital blindness); it excludes children who lost vision later in development and therefore had visual experience before suffering blindness, as well as children who have residual functional vision (visually impaired children).

Given the great individual differences that exist among congenitally blind children, it is difficult to depict a typical or characteristic portrait of these children, and generalizations must be considered with caution.

Blindness affects the development of congenitally blind children in different ways, and language development is one area minimally affected by the lack of vision.

The use and knowledge of language in blind children has been a controversial topic. The first studies on this population, carried out in the 1950s, 1960s,

and 1970s, were not very strict in terms of the method used. Nevertheless, they were influential in the scientific community and created among specialists the belief that blind children showed difficulties, delays, and malfunctions in their language development. Only quite recently did a different view on the acquisition of language by blind children appear in the field. This was so in part because in more recent studies, researchers applied modern techniques and methods to study the acquisition of language.

Meaning

Cutsford was the first scholar to claim that the words of blind children were meaningless, because the children tended to define words (such as *blood* and *moon*) using visual terms (such as *red* and *yellow*), which are not accessible to them. To use Cutsford's expression, there exists a "verbal unreality of words." However, there are problems with Cutsford's investigation. On one hand, conscious access to the meaning of words (metalinguistic knowledge) may be different from the real meaning that children have. On the other hand, Cutsford's conception of meaning is extremely limited. The meaning of a word does not come solely from sensory experience and its conceptualization (reference). Language itself is extremely important in the creation of meaning (sense relations and grammatical meaning). In addition, children learn the meanings of words by reading adults' intentions (pragmatic meaning) in contexts of significant social interactions in which the words are used repeatedly. Although studies did not support Cutsford's claim, the idea that blind children have problems in developing a conventional meaning for words persisted in the field.

Dunlea (1989) theorized that young blind children had a limited capacity for generalizing a given word for other items in the category and used a word (*dog*) for the original referent (a dog puppet) or only few items in the category. A limited cognitive-based theory of meaning underlies this consideration. In this view, concepts have a sensorial or referential basis. Because blind children have limited access to perceptual characteristics of the external world, their concepts show a limited nonabstract character and lack generalization. However, blind people are able to form concepts that are equivalent to those formed by sighted individuals. The point in question is that blind children, and particularly young blind children, may have difficulties in using a given word for a variety of items simply because they have restricted experience and cannot use, for instance, the word *dog* for a dog that is walking unless it barks, or a verb for an action performed by another person unless the action has an audible component. In any case, this limitation does not affect the ability of blind children to form the concepts underlying meanings, but only limits their access to external information.

The learning of the meanings of words for which blind children cannot have sensorial experience is critical to understanding the complexities of the process of meaning acquisition. Unexpectedly, blind children around 4 years of age are able to comprehend that words such as *green* or *red* refer to concrete characteristics of objects that they cannot perceive, although other sighted people can, and they may ask, "What color is this?" In a similar vein, blind children can distinguish between the meanings of *see* and *look* without having had visual experience. Landau and Gleitman (1985) showed how a blind girl (Kelly) knew that for somebody to see a given object, the object had to be in the viewer's line of sight, and there had to be no barrier between the person and the object. However, Kelly used *look* with the meaning of exploring objects with her hands.

What all these indicate is that blind children can use the information provided by language itself in order to learn the meanings of words. Color terms appear in the same position as other qualifying adjectives, and children can catch these positional regularities; that is, they can draw meaning from grammatical information as they acquire experience with language.

In summary, blind children use the same words as their sighted peers and attribute similar meanings to the words, although blind children probably rely on the information provided by language to a greater extent than sighted children.

Morphology

Most researchers have agreed that with the exception of personal and possessive pronouns, the morphological development of blind children is neither delayed nor impaired, although it is different from that of sighted children.

The results obtained by Dunlea and Andersen (1992) suggest that at first young blind children use few morphemes (plural, third person of present indicative, and locative prepositions *on* and *in*) in formulaic structures and imitations. Productive use of these morphemes appears later in blind children than in sighted children. In line with their view of language acquisition, these authors related this delay to the cognitive difficulties blind children have in conceptualizing the number of entities of the external world, the spatial relationships of entities, and the agency of others. However, first uses of morphemes in formulaic and limited-scope structures (pivot

schemas, item-based constructions, for example) are also typical of young sighted children; therefore, no special explanation would be necessary to understand blind children's language. In contrast, it seems that blind children use some morphemes, such as the regular past, earlier than sighted children, while other morphemes are used at similar ages by both groups. Blind children also use over-regularizations of irregular morphemes (*goed*, *singed*) at the same ages as their sighted peers. This indicates that verbal morphemes are the result of the use of a general rule, one that is drawn thanks to the linguistic analysis of input (Pérez-Pereira and Castro, 1997), contradicting the idea that blind children show a lack of creativity in their language.

Pronouns

According to the first studies on the topic (Fraiberg and Adelson, 1973; Dunlea, 1989), the use of personal and possessive pronouns and adjectives by blind children was deficient. In these studies, blind children started to use pronouns productively very late (nearly at age four), and they produced a great proportion of reversal errors (first-person for second-person pronouns and vice versa). These studies did not offer quantitative data, making replication impossible, and relied on authors' personal appreciation of data. However, the view they offered fueled the idea that blind children resemble autistic children, who typically produce numerous reversals. More recent research (Pérez-Pereira, 1999) found that not all blind children produced reversal errors and that most of them produced few errors of this type, similarly to what sighted children do. In addition, blind children's use of personal reference terms followed a developmental pattern similar to that of sighted children: they started to produce the first pronouns in formulaic expressions before 2 years of age, and between 2 and 3 years, all of them produced a great number and variety of pronouns in flexible structures. It was also found that imitation of a previous utterance produced by the interlocutor did not seem to be a sufficient explanation of reversals, contrary to one of the most widespread hypotheses. Children's ability to shift perspective and adopt a distinctive point of view seems to be necessary for correct production of personal reference terms. In any case, difficulties with pronouns are not a general characteristic of blind children as a group.

Syntax

As for syntactic development, compared with those of sighted children, MLU (mean length of utterance) scores of blind children throughout development indicate that blind children are not delayed.

When complexity of structures has been analyzed, blind children have shown development similar to that of sighted children not only during the first steps of grammatical development but also during the acquisition of complex sentences. Blind children's first use of coordinate and subordinate sentences starts at 2;6 years of age or shortly thereafter, and at 3;6, blind children produce a great variety of complex sentences, although they are not frequent. Around 4 years of age, blind children produce all types of coordinate and subordinate sentences.

Pragmatics

Blind children use language with similar functions, and learn to perform these functions, at the same ages as sighted children. Nevertheless, some differences exist up until 4;6 years; these are connected to the adaptive strategies that blind children put into practice, and/or to their limited access to information about external reality. However, these differences disappear with time (Pérez-Pereira and Castro, 1997). The main early difference is that blind children tend to use self-oriented language, instead of externally oriented language. Therefore, compared with sighted children, they use descriptions of their own actions or their own intentions to carry out an action, or expressions of their wishes, to a greater extent than descriptions of others' actions. Blind children produce language with reference to locations of objects and actions, qualities of objects, and the description of external events less frequently than their sighted peers, which is surely due to blind children's restricted access to information of the external world. Blind children also produce a lesser proportion of verbal expressions to offer, show, or draw another's attention (socially oriented speech) than sighted children, probably because they lack the motivating factor of perceiving their interlocutors.

Young blind children use a relatively high proportion of calls or vocatives as a way of getting information about the presence and location of other people in the immediate environment as well.

Another differential characteristic of blind children is their higher use of routines and repetitions/imitations in comparison to sighted children. Learning and using whole phrases or formulas for specific contexts and activities allows blind children to participate in social interactions and shared activities with other people (Peters, 1994). This finding is related to the strategy blind children follow in their learning of language.

Like children with autism, blind children have been described as having serious difficulties in

participating in communicative conversations (Moore and McConachie, 1994, among others). However, problems of method in many of these studies led the authors to reach wrong conclusions. When analyses were performed on the basis of frequency of initiations, and not proportion of initiations relative to mothers' initiations, it was found that blind children initiated as many conversations as nonimpaired children. In addition, they produced few conversational breakdowns and took as many turns as sighted children did, indicating that blind children and sighted children show similar abilities to participate in communicative conversations (Pérez-Pereira and Conti-Ramsdem, 1999, 2001, 2005).

Input

Andersen *et al.* (1993) said that apart from the conceptual consequences of being visually deprived, blind children generally receive less appropriate linguistic input than sighted children and that the combination of these two factors gives rise to their poor linguistic development. Contrary to previous studies (Moore and McConachie, 1994), the data we have today have indicated that the mothers of blind children use language well adapted to their needs (Pérez-Pereira and Conti-Ramsden, 1999, 2001, 2004). For instance, compared with the mothers of sighted children, mothers of blind children are more insistent in their intents to initiate and continue conversations with their children. In addition, the directives that mothers of blind children use include a much greater number of descriptions of objects, their locations, and the actions to be performed by their children.

Strategies

Convergent characteristics of the language they produce suggest that blind children are inclined to use a particular strategy or style of language acquisition: the rote-learned style or Gestalt style. Blind children tend to imitate a great number of formulas and frozen phrases first associated to certain scenes (Peters, 1987, 1994). These children tend to use frozen, or unanalyzed, phrases first and analyze them later. When blind children were compared with visually impaired and sighted children at the beginning of their combinatorial speech, it was found that blind children used many more frozen phases and that productive uses of a given combination appeared later in development and were used less frequently between 21 and 25 months of age (Pérez-Pereira, 2001, 2004). First combinations of words in blind children showed a clear transition from rote-learned phrases to item-based constructions and then toward a gradual process of schematization and generalization.

Blind children relied on the use of modified and expanded imitations and self-repetitions to a greater extent than did sighted children, and all children used these as a means to analyze language.

In general, the use of imitations and repetitions allows children to produce utterances that are morphologically more complex than their spontaneous utterances (Pérez-Pereira, 1994), which indicates the progressive role that imitations and repetitions play in grammar development.

Conclusion

The adoption of an individual-differences approach helps us to better understand the peculiarities of blind children's language acquisition. Former characterizations of blind children's language as 'parroting' or 'formulaic speech' should give way to an understanding of their language that takes into account the prevalence of a rote-learned or Gestalt style as a kind of early strategy, rather than as a deficiency.

The fact that congenitally blind children eventually manage to master language that does not differ from that of sighted people is a critical test for theories of language acquisition.

See also: Formulaic Speech; Language Development: Overview; Variation in First Language Acquisition.

Bibliography

Andersen E S, Dunlea A & Kekelis L S (1993). 'The impact of input: language acquisition in the visually impaired.' *First Language 13,* 23–49.

Cutsford T D (1951). *The blind in school and society.* New York: American Foundation for the Blind.

Dunlea A (1989). *Vision and the emergence of meaning: blind and sighted children's early language.* Cambridge, UK: Cambridge University Press.

Dunlea A & Andersen E S (1992). 'The emergence process: conceptual and linguistic influences on morphological development.' *First Language 12,* 95–115.

Fraiberg S & Adelson E (1973). 'Self-representation in language and play: observations of blind children.' *Psychoanalysis Quarterly 42,* 539–562.

Landau B & Gleitman L R (1985). *Language and experience: evidence from the blind.* Cambridge, MA: Harvard University Press.

Moore V & McConachie H (1994). 'Communication between blind children and severely visually impaired children and their parents.' *British Journal of Developmental Psychology 12,* 491–502.

Pérez-Pereira M (1994). 'Imitations, repetitions, routines, and the child's analysis of language: insights from the blind.' *Journal of Child Language 21,* 317–337.
Pérez-Pereira M (1999). 'Deixis, personal reference, and the use of pronouns in blind children.' *Journal of Child Language 26(3),* 655–680.
Pérez-Pereira M (2001). 'First grammar in blind, visually impaired and sighted bilingual children: do they follow a different route?' In Almgren A B M, Ezeizabarrena M J, Idiazabal I & MacWhinney B (eds.) *Research on child language acquisition. Proceedings of the VIIth International Congress for the Study of Child Language – IASCL.* Somerville, MA: Cascadilla Press. 1198–1208.
Pérez-Pereira M (2004). 'Diferencias individuales en el desarrollo gramatical inicial: aportaciones del estudio de niños ciegos.' *Anuario de Psicología 35(2),* 279–294.
Pérez-Pereira M & Castro J (1997). 'Language acquisition and the compensation of visual deficit: new comparative data on a controversial topic.' *British Journal of Developmental Psychology 15,* 439–459.
Pérez-Pereira M & Conti-Ramsden G (1999). *Language development and social interaction in blind children.* Hove, UK: Psychology Press.
Pérez-Pereira M & Conti-Ramsden G (2001). 'The use of directives in verbal interactions between blind children and their mothers.' *Journal of Visual Impairment and Blindness 95,* 133–149.
Pérez-Pereira M & Conti-Ramsden G (2004). Caracteristicas generales de las interacciones verbales entre niños ciegos y sus madres. *Infancia y Aprendizaje 26(3),* 381–396.
Pérez-Pereira M & Conti-Ramsden G (2005). 'Do blind children show autistic features?' In Pring L (ed.) *Autism and blindness: current findings and reflections.* London: Whurr Publishers. 99–127.
Peters A M (1987). 'The role of imitation in the developing syntax of a blind child.' *Text 7,* 289–311.
Peters A M (1994). 'The interdependence of social, cognitive, and linguistic development: evidence from a visually impaired child.' In Tager-Flusberg H (ed.) *Constraints on language acquisition: studies of atypical children.* Hillsdale, NJ: Erlbaum. 195–219.
Pring L (ed.) (2004). *Autism and blindness: current findings and reflections.* Londres: Whurr Publishers.
Warren D H (1994). *Blindness and children: an individual differences approach.* New York: Cambridge University Press.
Webster A & Roe J (1998). *Children with visual impairments: social interaction, language and learning.* New York: Routledge.

Language Development of Deaf Children with Hearing Parents

A Lederberg, Georgia State University, Atlanta, GA, USA

Deaf children who have deaf parents typically acquire language naturally through immersion in a sign language-rich environment. However, over 94% of deaf children are born with hearing parents and face a very different language learning situation than either hearing children or deaf children with deaf parents. These children are born into a world without language because they cannot perceive the spoken language in their environment. Initial exposure to language can occur any time from infancy through adulthood, and depends on both identification of hearing loss and placement in an appropriate intervention program. In general, intervention consists of attempting to make language accessible through one of three types of language learning environments or methodologies. First, oral interventions are designed to make spoken language accessible by use of amplification devices such as hearing aids or cochlear implants paired with intensive speech input. Until recently, for most deaf children, language skills remained very limited using this approach. However, since the 1980s, deaf children's ability to perceive speech has improved through the use of cochlear implants. These implants provide electrical stimulation directly to the endings of the auditory nerve and enable deaf children to perceive sound. While auditory input is still less rich and finely detailed than that received by hearing children, speech perception and spoken language development is typically enhanced by use of cochlear implants. A second common intervention is to communicate with both speech and sign simultaneously. The language in these simultaneous communication (SC) programs is that of the hearing culture (e.g., English, Italian, German) not that of natural sign languages like American Sign Language (ASL). For example, sign systems for English (collectively called manually coded English, or MCE) typically consist of borrowed lexical signs from ASL and invented signs that represent the functional aspects of English (e.g., articles, auxiliaries, copula, tense markers such as *-ed*, *-ing*) that are not expressed by isolated signs in ASL. These signs are then produced in English word order (Schick, 2003). Although SC programs are designed to make language 'visible' and thus fully accessible to deaf children, in reality, this rarely happens. Hearing parents and teachers vary greatly in their signing skills, and rarely sign all their

severe hearing impairment.' *Journal of Speech and Hearing Disorders 55,* 628–634.
Meadow-Orlans K P, Spencer P E & Koester L S (2004). *The world of deaf infants: a longitudinal study.* New York: Oxford University Press.
Moeller M P (2000). 'Early intervention and language development in children who are deaf and hard of hearing.' *Pediatrics 106(3),* e43 (41–49).
Moeller M P, Osberger M J & Eccarius M (1986). 'Receptive language skills.' *Language and Learning Skills in Hearing-Impaired Children 23,* 41–53.
Nicholas J G (2000). 'Age differences in the use of informative/heuristic communicative functions in young children with and without hearing loss who are learning spoken language.' *Journal of Speech, Language, and Hearing Research 43,* 380–394.
Nicholas J G & Geers A E (2005). 'The process and early outcomes of cochlear implantation by three years of age.' In Spencer P E & Marschark M (eds.) *Advances in the spoken language development of deaf children.* New York: Oxford University Press.
Oller D K & Eilers R E (1988). 'The role of audition in infant babbling.' *Child Development 59(2),* 441–449.
Osberger M J, Moeller M P, Eccarius M, Robins A M & Johnson D (1986). 'Expressive language skills.' *Monographs of the American Speech, Language, and Hearing Association 23,* 54–65.
Prezbindowski A K, Adamson L B & Lederberg A R (1998). 'Joint attention in deaf and hearing 22-month-old children and their hearing mothers.' *Journal of Applied Developmental Psychology 19(3),* 377–387.
Schauwers K, Gillis S & Govaerts P (2005). 'Language acquisition in children with a cochlear implant.' In Fletcher P & Miller J F (eds.), *Language disorders and developmental theory.* Philadelphia, PA: John Benjamins. 95–119.
Schick B (2003). 'The development of American Sign Language and manually coded English systems.' In Marschark M & Spencer P E (eds.) *Oxford handbook of deaf studies, language, and education.* New York: Oxford University Press. 219–231.
Spencer P E (1993). 'Communication behaviors of infants with hearing loss and their hearing mothers.' *Journal of Speech and Hearing Research 36,* 311–321.
Spencer P E (2004). 'Individual differences in language performance after cochlear implantation at one to three years of age: child, family, and linguistic factors.' *Journal of Deaf Studies and Deaf Education 9,* 393–412.
Stoel-Gammon C (1988). 'Prelinguistic vocalizations of hearing-impaired and normally hearing subjects: a comparison of consonantal inventories.' *Journal of Speech and Hearing Disorders 53,* 302–315.
Szagun G (2004). 'Learning by ear: on the acquisition of case and gender marking by German-speaking children with normal hearing and with cochlear implants.' *Journal of Child Language 31,* 1–30.
Taeschner T, Devescovi A & Volterra V (1988). 'Affixes and function words in the written language of deaf children.' *Applied Psycholinguistics 9,* 385–401.
Tait M, Lutman M E & Robinson K (2000). 'Preimplant measures of preverbal communicative behavior as predictors of cochlear implant outcomes in children.' *Ear and Hearing 21,* 18–24.
Tur-Kaspa H & Dromi E (1998). 'Spoken and written language assessment of orally trained children with hearing loss: syntactic structures and deviations.' *Volta Review 100,* 186–202.
Wallace V, Menn L & Yoshinaga-Itano C (2000). 'Is babble the gateway to speech for all children? A longitudinal study of children who are deaf or hard of hearing.' *The Volta Review 100(5),* 121–148.
Yoshinaga-Itano C, Sedey A L, Coulter D K & Mehl A L (1998). 'Language of early- and late-identified children with hearing loss.' *American Academy of Pediatrics 102(5),* 1161–1171.

Language Development in School-Age Children, Adolescents, and Adults

M A Nippold, University of Oregon, Eugene, OR, USA

In this article, language development beyond the preschool years is described, covering significant changes that occur in school-age children (ages 6–12 years), adolescents (ages 13–19 years), and adults (ages 20+ years). For many years, it was widely believed that language development was largely complete by the time a child entered school, at five or six years of age. More recently, however, it has become apparent that a great deal of linguistic growth occurs during the school-age years, adolescence, and well into adulthood, and that the development of language has no obvious point of completion (Nippold, 1998). This should not detract from the undeniable truth that language development in infants, toddlers, and preschool children is an amazing phenomenon, marked by rapid and salient change. By contrast, language growth beyond the preschool years is comparatively slow, subtle, and sometimes difficult to quantify. Nevertheless, later language development is an exciting area of scientific inquiry.

The differences between early and later language development are striking. For example, after steadily

progressing through the stages of cooing, babbling, and proto-words during the first year of life, infants emerge as toddlers who can make requests, call attention to themselves, and comment on the world around them using their first true words (*Hey! Doggie! Lookit!*). By their second birthday, children will have amassed a vocabulary of hundreds of different words, which they can use to express a wide variety of meanings in short simple utterances (*Billy run; Mommy buy book; Dolly pretty*), which soon will become long and complex sentences during the preschool years (ages 3–5) (Paul, 1981). By the time they reach their fifth birthday, children will be speaking intelligibly about topics beyond the 'here and now,' making few grammatical errors ("I wonder what we're going to do when we get to Grandma's house") (Leadholm and Miller, 1992), and drawing from a lexicon of over 15 000 different words. Upon entering kindergarten at age six years, children will demonstrate their linguistic prowess by engaging in extended conversations with family and friends, sharing personal anecdotes, retelling favorite bedtime stories, and convincing their parents to purchase a new kitten, scooter, or other desirable item, thereby giving the impression that language development is now complete (Nippold, 1998). Nevertheless, a great deal of language competence remains to be acquired. Factors that support and even promote the later development of language include metalinguistic, cognitive, and social development; the newly acquired ability to read; and growth in world knowledge and experience.

Lexical Development

By six years of age, most children have acquired a vocabulary of approximately 18 000 different words, many of which are the names of concrete objects (*bicycle, skates*) and common actions (*jump, catch*). At this young age, children's lexicons will also include many adjectives (*sharp, noisy*) and adverbs (*quickly, slowly*) that are used to modify a burgeoning repertoire of nouns and verbs. Most of these words will have been learned through informal spoken-language interactions with adults and other children, listening to conversations and stories, and having the opportunity to talk about events that occur in children's daily lives. Then, once children are able to read proficiently, usually by around age nine or 10 years, a new source of vocabulary learning – written language – becomes available to augment these spoken-language opportunities (Sternberg, 1987). With this newly acquired ability to consume a variety of books, magazines, and newspapers, children now gain exposure to a large number of difficult words that occur infrequently in spoken-language contexts (Cunningham and Stanovich, 1998; Nagy *et al.*, 1985; Stanovich and Cunningham, 1992). As a result, their lexicons will gradually expand during the preadolescent (ages 9–12) and adolescent (ages 13–19) years so that by early adulthood, they will understand and use at least 60 000 different words, including a large number of abstract nouns (*courage, freedom*), later-developing metalinguistic (*persuade, convince*) and metacognitive (*deduce, discover*) verbs, and other terms used to express precise meanings such as adverbs of magnitude (*decidedly, slightly*), adjectives of emotion (*delirious, effervescent*), adverbial conjuncts (*moreover, nevertheless*), and specialized terminology used in the study of academic subjects such as mathematics (*concentric, hyperbola*), biology (*monozygotic, phylogeny*), and history (*dictatorship, liberation*) (Nippold, 1998). Indeed, this capacity for acquiring new words and for expanding one's knowledge of old ones continues throughout the lifespan, particularly in individuals who are active readers, writers, speakers, and acquired who enjoy communicating with others.

Many of these new words will be acquired through metalinguistic strategies that older children, adolescents, and adults have learned to employ when reading. These include inferring the meaning of a word from the broader linguistic context in which it occurs, performing a morphological analysis of the word itself, and combining these strategies, as necessary (Nippold, 1998). For example, a general understanding of the word *dictatorship* could be inferred from reading a passage about Mussolini in a world history book. The learning process could be enhanced when the reader carefully considers the individual morphemes comprising this word, which include the lexical morpheme *dictate* 'to give orders to other people,' and the derivational morphemes *-or* 'a person who performs some action' and *-ship* 'a state of existence' to gain a more complete understanding of the word (i.e., a state of existence where one person rules with absolute power) (Nippold *et al.*, 2005a).

Metalinguistic competence is also reflected in the young person's growing ability to define increasingly difficult words using the formal 'Aristotelian' style. Frequently modeled in educational settings, this manner of explicating the meaning of a word involves stating the category to which it belongs and one or more key features, i.e., *X is a Y that Z*. By 12 years of age, most children can define concrete nouns in this way (e.g., *An* ***emu*** *is a bird that lives in Australia, is very large, but does not fly*), but it is not until late adolescence or even early adulthood that individuals can define abstract nouns in this concise, efficient, and informative manner (e.g., ***Idleness*** *is a*

state of inactivity that can include a lack of motivation or opportunity) (Nippold *et al.*, 1999). An important skill in literate language contexts such as formal lectures and reports, word definition can reveal the individual's depth and breadth of knowledge as key terminology is explained. Thus, it is not surprising that word definition is closely associated with academic achievement in children and adolescents (Nippold, 1998).

Semantic Development

An important aspect of semantics that gradually develops during the school-age years, adolescence, and adulthood is the understanding of figurative expressions such as 'metaphors' (*The bird was a flying rainbow*; *Genius is perseverance in action*), 'idioms' (*talk through one's hat; a wolf in sheep's clothing*), and 'proverbs' (*Every cloud has a silver lining; Humility often gains more than pride*). Although comprehension of these expressions begins in early childhood, the learning process continues throughout the lifespan, a situation that is evident when people are asked to interpret low-familiarity expressions that contain abstract vocabulary and are meant to convey complex thoughts.

As with new vocabulary learning, figurative expressions are learned through metalinguistic strategies. One such strategy, 'contextual abstraction,' occurs when the learner attempts to infer meaning from the larger linguistic context in which the expression is used. An additional strategy, 'metalexical analysis,' occurs when the learner performs an internal analysis of the expression to infer meaning from its individual words. Additionally, the amount of exposure an individual receives to specific expressions in meaningful communicative contexts makes a significant contribution to the learning process, thereby demonstrating the role of 'world experience.'

Metaphors

Many six-year-old children can comprehend simple perceptual metaphors such as *The bird was a flying rainbow*, expressions that refer to concrete entities that can easily be visualized (Nippold *et al.*, 1984). To accomplish this, children are able to reason by analogy to discern the relevant overlapping features between the topic (*bird*) and the vehicle (*rainbow*) of the metaphor to determine the common ground (*colorfulness*). What develops, as children grow older, is their ability to comprehend metaphors that contain complex vocabulary, that speak to human emotions, and that require a greater degree of world experience to be fully appreciated (Nippold, 1998). Examples of more difficult expressions include *Suspicion is quicksand*, *Trust is a relationship's glue*, and *Your call was a winged car*, psychological metaphors that may challenge the interpretative abilities of adolescents and young adults. When asked to interpret these sorts of metaphors, elderly adults often provide more insightful responses than younger individuals because of their extensive life experiences (Boswell, 1979).

Idioms

As with metaphors, idioms can vary widely in their degree of difficulty. Easier idioms are those that are more transparent and frequently occur in spoken and written language (Nippold and Rudzinski, 1993; Nippold *et al.*, 1996). For example, the figurative meaning of *go by the book*, to be highly compliant, is a metaphoric extension of its literal meaning, to read a document carefully. Not surprisingly, many preadolescent children readily understand this figurative expression. In contrast, more challenging idioms include those that are opaque and rarely occur in the language. For example, the figurative meaning of *go by the board*, that something has been discarded, is difficult to discern from a literal interpretation, even for many adults, particularly if they are unfamiliar with customs associated with sailing ships! Another factor that affects the ease with which idioms can be understood is the learner's ability to infer meaning from the broader linguistic context in which the expression occurs (Nippold and Martin, 1989), a skill that is related more generally to listening and reading comprehension (Nippold *et al.*, 2001). As with metaphors, older adults often provide more insightful interpretations of idioms (*take someone under one's wing*) than school-age children, adolescents, and younger adults (Nippold and Duthie, 2003).

Proverbs

Proverbs (*Scalded cats fear even cold water; Envy is destroyed by true friendship*) are expressions that capture the shared beliefs or collective wisdom of a society. For this reason, they are considered to be an important part of cultural literacy. Proverbs typically are more difficult to comprehend than metaphors and idioms because of the greater degree of world experience they require, their lower frequency of occurrence in the language, and the greater amount of verbal reasoning they entail (Nippold, 2000). Proverb learning occurs during the interplay between top-down pragmatic and bottom-up semantic processing strategies. By attending to the linguistic context in which an expression occurs, the learner gains information about why the speaker used the proverb. Then, by focusing on the individual words comprising the expression, the learner begins to draw analogies from the proverb itself to the larger context. For

example, a parent might use the proverb *scalded cats fear even cold water* to comment on her young son's refusal to try new foods, having recently experienced food poisoning. An older sibling unfamiliar with this proverb might begin to glean the metaphorical relationship that is being expressed between cats that have been scalded and the brother who was sickened, realizing that both are reacting negatively to anything related to their unfortunate experience. Thus, to understand a proverb, it is not sufficient to have knowledge of the words it contains. One also must draw analogies between items stated in the proverb and those existing in the environment, a process called 'metaphorical mapping' (Nippold *et al.*, 2000; Nippold and Haq, 1996; Nippold *et al.*, 1988).

Syntactic Development

Compared to the rapid and highly salient changes that occur during the preschool years, later syntactic development proceeds at a more modest pace, characterized by the production of increasingly longer and more complex sentences in spoken and written communication (Loban, 1976). For example, in a sample of conversational speech, a 6-year-old child will produce sentences of approximately six words in length, on average, but a 30-year-old adult will exhibit a mean sentence length of at least 10 words.

In addition, the use of sophisticated grammatical constructions and low-frequency syntactic structures gradually increases during the developmental period between 6 and 30 years (Berman and Verhoeven, 2002; Nippold *et al.*, in press; Scott, 1988; Verhoeven *et al.*, 2002). Key grammatical constructions include relative clauses (*You want to keep the ball in this area* ***that's called the fairway***), adverbial clauses (***If you get it in fewer strokes than par,*** *that's a birdie*), and nominal clauses (*There's an area beyond the fairway* ***where the grass is longer***); key low-frequency structures include noun phrase postmodification via prepositional phrase (*Birds* ***of the forest*** *build nests in diverse locations*), the passive voice (*The library* ***was established by a wealthy community member***), appositives (*Jason,* ***the student body president,*** *plays basketball*), the perfect aspect (*Amanda* ***had been working*** *all day*), compound auxiliary verbs (*We* ***might have*** *wanted to go hiking*), and nonfinite verb phrases such as participials (***Jumping into the lake,*** *Timmy held his breath*), gerunds (***Camping at the lake*** *is his favorite activity*), and infinitives (***To win the lottery*** *would be amazing*).

Although many of these syntactic devices are first employed during the preschool years, older children and adolescents are learning to use them in a more fluent and efficient manner to produce larger chunks of discourse. This includes, for example, engaging in lengthy conversations, telling and retelling favorite stories and personal anecdotes, and producing elaborate explanations of how things work. Fueling this growth in syntactic development is the individual's expanding knowledge base and increasingly complex thoughts (Nippold *et al.*, in press). When school-age children, adolescents, and adults are engaged in expository discourse (e.g., explaining how to play their favorite game or sport), they typically produce longer sentences containing greater amounts of subordination compared to when they are engaged in simple conversations about topics that have been discussed many times before (e.g., pets, siblings, or school).

Expository discourse, with its heightened emphasis on accuracy and detail in the explication of information, causes speakers to tap into their own cognitive and linguistic resources more fully than simpler types of discourse, revealing a greater degree of syntactic competence. For example, in a conversation about school, a 17-year-old girl produced a mean utterance length of 9.45 words. However, when she explained specific strategies for winning a card game, her mean utterance length grew to 13.65 words, and her use of all types of subordinate clauses increased, evident in the following utterances (From the author's files. Clause types, marked at the end of each clause, are as follows: IC = independent clause; AVC = adverbial clause; RC = relative clause; NOM = nominal clause):

(1) If you have more than one wild card [AVC], only put one down at a time [IC] until the pile that you have to pick up from [RC] is gone [AVC]. And when you're putting three cards on top [AVC] that are face up [RC], you probably want to make them like wild and high cards [IC], so you can get rid of them easier [NOM]. And if you have the same kind [AVC], keep them both there [IC] so you can get rid of them sooner [NOM].

Pragmatic Development

During the school-age years, adolescence, and into adulthood, pragmatic development is evident in a variety of communicative genres, including conversational, narrative, and persuasive discourse. Supported by growth in social and cognitive arenas, change is particularly obvious during adolescence when peer interactions become an increasingly important part of daily life, and the youth begins to reason more flexibly.

During the years between late childhood and early adulthood, the amount of time that young people spend socializing increases dramatically along with the number of hours they spend in conversation with friends, either in person or over the telephone

(Raffaelli and Duckett, 1989; Rawlins, 1992). Time spent socializing with friends provides the speaker the opportunity to practice new skills and to receive honest feedback. Conversational skills that are known to improve during the adolescent years include staying on topic, making relevant comments, interrupting appropriately, contributing novel ideas, making smooth transitions between topics, showing empathy, and entertaining the listener by telling jokes, offering amusing anecdotes, and emphasizing key points through dramatic gestures and facial expressions (Nippold, 1998, 2000).

The ability to adjust the tone and content of one's discourse also improves during these years, with adolescents and young adults showing a greater awareness of other people's thoughts, feelings, and emotions than children. This competence in perspective taking is exhibited in all realms of social communication including narrative and persuasive discourse. For example, in narrative discourse, older individuals are more likely to offer details concerning the inner thoughts, struggles, and reactions of the characters in a story, and in persuasive discourse, they are more likely to structure an argument that will appeal to the listener's values, beliefs, and priorities. Older speakers also can provide a greater number of arguments to support their own views (Nippold, 1998). At the same time, however, adolescents and adults are more likely than children to consider both sides of a controversy, and to evidence greater flexibility of thought in reaching balanced solutions (Nippold *et al.*, 2005b).

Conclusions

In recent years, later language development has become a topic of expanding international interest, with a rapidly growing number of articles, books, and conference presentations being devoted to this theme. Language development beyond the preschool years is an exciting and dynamic field of study, and one that is being fueled by the collaborative efforts of scientists around the world (Berman, 2004). In the future, it is expected that numerous crosslinguistic studies of later language development will be conducted, along with studies that closely examine languages other than English. These efforts undoubtedly will contribute substantially to the knowledge base in this area.

See also: Classroom Talk; Discourse, Narrative and Pragmatic Development; Family Speak; Figures of Speech; Influence of Literacy on Language Development; Language Development: Overview; Language Education: Vocabulary; Metaphor: Psychological Aspects; Metapragmatics.

Bibliography

Berman R A (2004). 'Between emergence and mastery: the long developmental route of language acquisition.' In Berman R A (ed.) *Language development across childhood and adolescence* (vol. 3). *Trends in language acquisition research (TILAR)*. Amsterdam: John Benjamins. 9–34.

Berman R A & Verhoeven L (2002). 'Cross-linguistic perspectives on the development of text-production abilities: speech and writing.' *Written Language and Literacy 5(1),* 1–43.

Boswell D A (1979). 'Metaphoric processing in the mature years.' *Human Development 22,* 373–384.

Cunningham A E & Stanovich K E (1998). 'What reading does for the mind.' *American Educator 22(1),* 8–15.

Leadholm B J & Miller J F (1992). *Language sample analysis: the Wisconsin guide.* Madison, WI: Wisconsin Department of Public Instruction.

Loban W (1976). *Language development: kindergarten through grade twelve. (Research Report No. 18.)* Urbana, IL: National Council of Teachers of English.

Nagy W E, Herman P A & Anderson R (1985). 'Learning words from context.' *Reading Research Quarterly 20,* 233–253.

Nippold M A (1998). *Later language development: the school-age and adolescent years* (2nd edn.). Austin, TX: Pro-Ed.

Nippold M A (2000). 'Language development during the adolescent years: Aspects of pragmatics, syntax, and semantics.' *Topics in Language Disorders 20(2),* 15–28.

Nippold M A, Allen M M & Kirsch D I (2000). 'How adolescents comprehend unfamiliar proverbs: the role of top-down and bottom-up processes.' *Journal of Speech, Language, and Hearing Research 43,* 621–630.

Nippold M A & Duthie J K (2003). 'Mental imagery and idiom comprehension: a comparison of school-age children and adults.' *Journal of Speech, Language, and Hearing Research 46,* 788–799.

Nippold M A, Duthie J K & Larsen J (2005a). 'Literacy as a leisure activity: free-time preferences of older children and young adolescents.' *Language, Speech, and Hearing Services in Schools 36*, 93–102.

Nippold M A & Haq F S (1996). 'Proverb comprehension in youth: the role of concreteness and familiarity.' *Journal of Speech and Hearing Research 39,* 166–176.

Nippold M A, Hegel S L, Sohlberg M M & Schwarz I E (1999). 'Defining abstract entities: development in pre-adolescents, adolescents, and young adults.' *Journal of Speech, Language, and Hearing Research 41,* 473–481.

Nippold M A, Hesketh L J, Duthie J K & Mansfield T C (in press). Conversational versus expository discourse: a study of syntactic development in children, adolescents, and adults. *Journal of Speech, Language and Hearing Research*.

Nippold M A, Leonard L B & Kail R (1984). 'Syntactic and conceptual factors in children's understanding of metaphors.' *Journal of Speech and Hearing Research 27,* 197–205.
Nippold M A & Martin S T (1989). 'Idiom interpretation in isolation versus context: a developmental study with adolescents.' *Journal of Speech and Hearing Research 32,* 59–66.
Nippold M A, Martin S A & Erskine B J (1988). 'Proverb comprehension in context: a developmental study with children and adolescents.' *Journal of Speech and Hearing Research 31,* 19–28.
Nippold M A, Moran C & Schwarz I E (2001). 'Idiom understanding in pre-adolescents: synergy in action.' *American Journal of Speech–Language Pathology 10,* 169–179.
Nippold M A & Rudzinski M (1993). 'Familiarity and transparency in idiom explanation: a developmental study of children and adolescents.' *Journal of Speech and Hearing Research 36,* 728–737.
Nippold M A, Taylor C L & Baker J M (1996). 'Idiom understanding in Australian youth: a cross-cultural comparison.' *Journal of Speech and Hearing Research 39,* 442–447.
Nippold M A, Ward-Lonergan J M & Fanning J L (2005b). 'Persuasive writing in children, adolescents, and adults: a study of syntactic, semantic, and pragmatic development.' *Language, Speech, and Hearing Services in Schools 36,* 125–138.
Paul R (1981). 'Analyzing complex sentence development.' In Miller J F (ed.) *Assessing language production in children: experimental procedures.* Baltimore, MD: University Park Press. 36–40.
Raffaelli M & Duckett E (1989). '"We were just talking . . .": conversations in early adolescence.' *Journal of Youth and Adolescence 18,* 567–582.
Rawlins W K (1992). *Friendship matters: communication, dialectics, and the life course.* NY: DeGruyter.
Scott C M (1988). 'Spoken and written syntax.' In Nippold M A (ed.) *Later language development: ages nine through nineteen.* Boston: College-Hill. 49–95.
Stanovich K E & Cunningham A E (1992). 'Studying the consequences of literacy within a literate society: the cognitive correlates of print exposure.' *Memory and Cognition 20(1),* 51–68.
Sternberg R J (1987). 'Most vocabulary is learned from context.' In McKeown M G & Curtis M E (eds.) *The nature of vocabulary acquisition.* Hillsdale, NJ: Erlbaum. 89–105.
Verhoeven L, Aparici M, Cahana-Amitay D, van Hell J, Kriz S & Viguie-Simon A (2002). 'Clause packaging in writing and speech: a cross-linguistic developmental analysis.' *Written Language and Literacy 5(2),* 135–162.

Language Development: Morphology

H Behrens, Rijksuniversiteit, Groningen, The Netherlands

Most studies on the acquisition of morphology focus on inflectional morphology (but see Clark, 1993 for results on compounding and derivation). Of particular interest are studies on the typical order of acquisition of morphology, as well as on individual and crosslinguistic differences and the causes underlying such variation.

In many languages, the earliest stages of language development show characteristics of 'telegraphic speech.' That is, inflectional morphemes and function words are initially absent but are gradually added on in the process of acquisition.

Studies of Morpheme Order Acquisition

Brown (1973: 271, Figure 14) provided a thorough methodological milestone for the systematic comparison of individual developmental trajectories. Brown related the acquisition of 14 carefully selected grammatical morphemes to stages of development, as defined by the mean length of utterance (MLU), and productivity, as defined by 90% of use in obligatory contexts. Upon a systematic scoring, the three English-speaking children under investigation showed a high correlation in their acquisition of the 14 morphemes. Present progressive was typically the first morpheme to become productive, followed by the prepositions *in/on*, plural and past irregular marking (Brown, 1973: 274, Table 38). Brown's results corresponded to the results from previous diary studies and experimental data (Brown, 1973: 275ff.) and inspired a wide range of comparative research.

Individual Differences

The availability of more and more data on individual developmental trajectories revealed not only that children differ in the age at which they acquire particular morphemes, but also that children show qualitatively different styles of language acquisition regarding the units they first pick up from the input sound stream. While most children initially identify stressed stems without unstressed inflectional elements, other

Pérez-Pereira M (1994). 'Imitations, repetitions, routines, and the child's analysis of language: insights from the blind.' *Journal of Child Language 21,* 317–337.
Pérez-Pereira M (1999). 'Deixis, personal reference, and the use of pronouns in blind children.' *Journal of Child Language 26(3),* 655–680.
Pérez-Pereira M (2001). 'First grammar in blind, visually impaired and sighted bilingual children: do they follow a different route?' In Almgren A B M, Ezeizabarrena M J, Idiazabal I & MacWhinney B (eds.) *Research on child language acquisition. Proceedings of the VIIth International Congress for the Study of Child Language – IASCL.* Somerville, MA: Cascadilla Press. 1198–1208.
Pérez-Pereira M (2004). 'Diferencias individuales en el desarrollo gramatical inicial: aportaciones del estudio de niños ciegos.' *Anuario de Psicología 35(2),* 279–294.
Pérez-Pereira M & Castro J (1997). 'Language acquisition and the compensation of visual deficit: new comparative data on a controversial topic.' *British Journal of Developmental Psychology 15,* 439–459.
Pérez-Pereira M & Conti-Ramsden G (1999). *Language development and social interaction in blind children.* Hove, UK: Psychology Press.
Pérez-Pereira M & Conti-Ramsden G (2001). 'The use of directives in verbal interactions between blind children and their mothers.' *Journal of Visual Impairment and Blindness 95,* 133–149.
Pérez-Pereira M & Conti-Ramsden G (2004). Caracteristicas generales de las interacciones verbales entre niños ciegos y sus madres. *Infancia y Aprendizaje 26(3),* 381–396.
Pérez-Pereira M & Conti-Ramsden G (2005). 'Do blind children show autistic features?' In Pring L (ed.) *Autism and blindness: current findings and reflections.* London: Whurr Publishers. 99–127.
Peters A M (1987). 'The role of imitation in the developing syntax of a blind child.' *Text 7,* 289–311.
Peters A M (1994). 'The interdependence of social, cognitive, and linguistic development: evidence from a visually impaired child.' In Tager-Flusberg H (ed.) *Constraints on language acquisition: studies of atypical children.* Hillsdale, NJ: Erlbaum. 195–219.
Pring L (ed.) (2004). *Autism and blindness: current findings and reflections.* Londres: Whurr Publishers.
Warren D H (1994). *Blindness and children: an individual differences approach.* New York: Cambridge University Press.
Webster A & Roe J (1998). *Children with visual impairments: social interaction, language and learning.* New York: Routledge.

Language Development of Deaf Children with Hearing Parents

A Lederberg, Georgia State University, Atlanta, GA, USA

Deaf children who have deaf parents typically acquire language naturally through immersion in a sign language-rich environment. However, over 94% of deaf children are born with hearing parents and face a very different language learning situation than either hearing children or deaf children with deaf parents. These children are born into a world without language because they cannot perceive the spoken language in their environment. Initial exposure to language can occur any time from infancy through adulthood, and depends on both identification of hearing loss and placement in an appropriate intervention program. In general, intervention consists of attempting to make language accessible through one of three types of language learning environments or methodologies. First, oral interventions are designed to make spoken language accessible by use of amplification devices such as hearing aids or cochlear implants paired with intensive speech input. Until recently, for most deaf children, language skills remained very limited using this approach. However, since the 1980s, deaf children's ability to perceive speech has improved through the use of cochlear implants. These implants provide electrical stimulation directly to the endings of the auditory nerve and enable deaf children to perceive sound. While auditory input is still less rich and finely detailed than that received by hearing children, speech perception and spoken language development is typically enhanced by use of cochlear implants. A second common intervention is to communicate with both speech and sign simultaneously. The language in these simultaneous communication (SC) programs is that of the hearing culture (e.g., English, Italian, German) not that of natural sign languages like American Sign Language (ASL). For example, sign systems for English (collectively called manually coded English, or MCE) typically consist of borrowed lexical signs from ASL and invented signs that represent the functional aspects of English (e.g., articles, auxiliaries, copula, tense markers such as *-ed*, *-ing*) that are not expressed by isolated signs in ASL. These signs are then produced in English word order (Schick, 2003). Although SC programs are designed to make language 'visible' and thus fully accessible to deaf children, in reality, this rarely happens. Hearing parents and teachers vary greatly in their signing skills, and rarely sign all their

spoken words. Therefore, the degree to which signed input represents language is variable. The third type of intervention, bilingual programs, use a natural sign language (e.g., ASL) as their mode of communication and typically expose the children to spoken language solely through the written medium.

This chapter focuses on deaf children's acquisition of the language of the hearing culture, whether through oral or SC methods. In addition to variations in language input, characterizing the language development of deaf children is further complicated by the fact that different cohorts of deaf children have experienced different language learning environments (Marschark, 2001). During the last decade, universal newborn screening for hearing loss has been instituted in many countries; and deaf children are beginning to gain access to language during infancy. For any aspect of language that is subject to a critical period, these changes will result in different developmental pathways (Lederberg and Spencer, 2005; Yoshinaga-Itano *et al.*, 1998). In addition, because of earlier identification and changing government policy, the age that children can receive cochlear implants has fallen from preschool to infancy over the last 5 years. Finally, improvements in both hearing aids and cochlear implants are ongoing, resulting in different auditory experiences for different cohorts of children (Nicholas and Geers, 2005; Schauwers *et al.*, 2005).

In comparison to hearing children, deaf children's language and communication development shows three types of patterns, which will be highlighted in this review. First, some robust phenomena develop identically for deaf and hearing children. Other aspects are delayed but exhibit similar developmental pathways. In these cases, deaf children resemble younger language-matched hearing children. Finally, some aspects are qualitatively different, with characteristics that deviate from that found among hearing children.

Babbling

Hearing infants begin to produce well-formed speech-like consonant-vowel reduplicated syllables (mamama, dadadada), or canonical babbles, around 6–10 months. Canonical babbles do not appear in deaf infants' vocalizations until 12–25 months, and account for a much lower ratio of their vocalizations (Oller and Eilers, 1988). In addition, the size and diversity of the consonant inventory in deaf children's babbles is greatly reduced (Stoel-Gammon, 1988). Cochlear implantation results in the development of canonical babbling that more closely resembles that of hearing toddlers with the onset of babbling occurring after only a few months (range 1–6.5 months) of auditory experience (Schauwers *et al.*, 2005). Although delayed in terms of chronological age because cochlear implantation does not occur at birth, this is one of the few areas that rate of development seems to be accelerated in terms of hearing age (i.e., length of time since cochlear implantation).

The relation between babbling and later language development is far from clear. On the one hand, the quality of children's canonical babbles seems to be an indicator of phonological knowledge (Oller and Eilers, 1988). However, babbling may be neither predictive nor necessary for later language abilities. Wallace *et al.* (2000) found that the quality of speech in elementary school correlated with variety of consonants in vocalizations during preschool but not the amount of canonical babbles in infancy. For children with cochlear implants, accelerated onset of babbling does not necessarily result in accelerated language development.

Pragmatics

Preschool Communication

While early linguistic development is affected by deafness, early **communication** development is remarkably similar for deaf and hearing toddlers. Both deaf and hearing toddlers' engagement in coordinated joint attention with their hearing mothers begins to appear at 12 months and shows similar, dramatic increases throughout the 2nd year of life (Meadow-Orlans *et al.*, 2004; Prezbindowski *et al.*, 1998). Hearing and deaf toddlers intentionally communicate with equal frequency, communicating primarily through vocalizations (Lederberg and Everhart, 2000; Spencer, 1993). However, because of a lack of linguistic intonational markers, the pragmatic functions of deaf children's vocalizations are more difficult to interpret. Therefore, they might be expected to use gestures more frequently than hearing toddlers. However, research shows deaf and hearing toddlers use similar gestural communications at 12–22 months old (Lederberg and Everhart, 2000; Spencer, 1993). Even more surprisingly, nonlinguistic gestural communication during this early period is similar for even deaf toddlers with deaf parents, even though these children are exposed to more gestures in their linguistic environment (Schick, 2003). Thus, early gesture development seems unaffected by either hearing loss or large linguistic environmental variations.

However, as deaf children become increasingly language delayed, there are marked differences in their communication skills. By 22 months, deaf children communicate less frequently and spend less time than

hearing children in joint attention with their hearing mothers, probably because joint attention becomes enriched and extended for the hearing children as language is integrated into their attentional episodes (Prezbindowski *et al.*, 1998). Pragmatic functions are also more ambiguous and different than those of hearing children (Lederberg and Everhart, 2000; Nicholas, 2000). Deaf toddlers are more likely to direct their mothers and less likely to make statements or ask questions than hearing children. During preschool, their use of directives decreases and statements increases. However, question-asking (even through nonverbal means) remains infrequent. The pragmatic function of deaf children's early communication, especially question-asking, is related to their language abilities (both concurrent and predictive) (Lederberg and Everhart, 2000; Nicholas, 2000). For children who receive a cochlear implant, the amount their prelinguistic communication is self-initiated (i.e., not solely a response to their mothers' communication) correlates with later language development after cochlear implantation (Tait *et al.*, 2000). While nonverbal communication is sufficient to make requests and direct attention, information exchange (comments and question-asking) may depend on, and perhaps is a motivator for, language development.

Referential Communication

Older deaf children continue to show deficits in their pragmatic skills. The ways that deaf children and adolescents seek information in a standard referential communication task are immature and relatively ineffective. Jeanes *et al.* (2000) found that both oral and SC deaf students (8- to 17-year-olds) were more likely to ask for clarifications using general requests (e.g., tell me again) than specific requests (e.g., Which color is it?), the latter being the most common repair among hearing dyads. In addition, signing deaf children's responses to requests for clarification were more likely to be merely repetitions of the original communication and less likely to contain new information compared to repairs by hearing and oral deaf children. Even during adolescence, conversational repairs by deaf students were similar to young hearing children rather than their agemates.

Narratives

Deaf children in oral and SC environments (with and without cochlear implants) show deficits in their narrative skill (Crosson and Geers, 2001). Compared to age-matched hearing children, deaf children spend more time describing the setting and less time providing a resolution to the story. They are less likely to link their propositions with subordinating, coordinating, and temporal conjunctions than hearing children. Thus, their stories show less cohesion. Although poor syntactic knowledge may cause some of the problems in narrative skill, immature story development may also be due to children's lack of exposure to stories because of poor reading abilities and an education focused on basic skill development. In fact, some subgroups of deaf children do have near age-appropriate skills. Crosson and Geers (2001) found that, unlike children with poor speech perception skills, those deaf children with cochlear implants who had above average speech perception skills told stories that were similar in narrative structure (i.e., containing orientations, complicating actions, evaluations, and resolutions) and reference as hearing children. Marschark *et al.* (1994) found no differences in the causal network structure (i.e., goal-action-outcome episodes) of stories generated by hearing children and those generated by SC deaf children and adolescents from a residential school for the deaf. Both sets of deaf children may have a rich language environment that exposes them to the classic English narrative structure.

Vocabulary

Lexical Growth

Past research has shown that deaf children have small lexicons, averaging 2 standard deviations below hearing children and show a rate of acquisition of words at 40–60% of that of hearing children, even when the children are provided with consistent amplification or cochlear implants and high-quality programming (Lederberg, 2003). Some deaf children with vocabularies of over 100 words are still restricted to slow word learning, adding only a few new words to their lexicon in a month, something never observed among hearing children. Even more disturbing, in a cross-sectional study, lexicon size seemed to plateau around 12 years of age, with deaf children 13 through 20 years old remaining at an age equivalent of hearing 9-year-olds on the Peabody Picture Vocabulary Test (PPVT) (Moeller *et al.*, 1986).

Recent studies have found that children who were identified in infancy and received appropriate intervention (which sometimes included cochlear implants) experienced faster vocabulary growth than in previous research, although development was still highly variable. In a study of vocabulary development of over 200 children (none with cochlear implants) enrolled in a state-wide early intervention program, Mayne *et al.* (2000) found that deaf toddlers with no cognitive disabilities produced, on

average, about 45 words at two years of age, and 350 words at 3 years. Vocabulary development was affected by age of identification (before vs. after 6 months) and cognitive impairments, but not communication method (sign vs. speech), degree of hearing loss, gender, ethnicity, or education of caregiver. Close to age-appropriate lexical growth is possible for young deaf children. On average, lexicon size of children who were identified before 6 months and were not cognitively impaired was comparable to age-matched hearing children in the 5th to 25th percentile. On the other hand, the lexicons of children who were in the lower percentiles of deaf norms were much smaller than any observed in hearing children.

Research by Moeller (2000) indicates that age-typical vocabulary knowledge can occur for some deaf children throughout preschool. Half of 112 deaf 5-year-olds who were enrolled in early intervention programs scored within 1 standard deviation of PPVT hearing norms. Age of intervention (before and after 11 months) and family involvement were the only factors affecting lexical knowledge. While children with highly involved parents scored close to hearing norms, children with parents who were involved at a moderate or low level had standard scores ranging from 60 to 80 depending on their age of identification. Similar to Mayne *et al.*'s research (2000), communication mode did not affect lexicon size. This is likely because children were placed in language environments (oral vs. SC) according to their needs.

There is a similar range of individual differences in lexical development among children with cochlear implantation (Nicholas and Geers, 2005; Schauwers *et al.*, 2005). However, overall, children with cochlear implants experience faster vocabulary growth than children with traditional hearing aids. In addition, although the research is not unequivocal, research suggests there may be a critical period in the effectiveness of cochlear implants (Lederberg and Spencer, 2005; Nicholas and Geers, 2005). Implantation before 2 years of age appears to result in deaf children acquiring spoken words at a rate near that expected for hearing children, with rates declining as age of implantation increases.

Word-Learning Processes

Given delayed and slower vocabulary development, the development of processes used to establish meaning is remarkably resilient among deaf children. Lederberg and colleagues examined two levels of fast-mapping abilities in 100 deaf preschoolers (Lederberg *et al.*, 2000; Lederberg and Spencer, 2005). Both cross-sectional and longitudinal research indicated that children's ability to fast-map (defined as quickly learning the meaning of) new words was primarily associated with lexicon size, rather than age. Children who were in the initial stages of vocabulary development (as measured by the number of words they produced) did not fast-map words. Children with moderately sized lexicons learned to understand the meaning of new words quickly but only when the meaning of a word was made explicit by social and pragmatic cues. Finally, those children with larger lexicons were also able to infer the meaning of words by the context (i.e., that a novel word refers to a novel object). Relations between lexicon size and word learning processes were the same for children acquiring spoken or signed English, those with a cochlear implant and those without, and for children who ranged in age of exposure to accessible input from birth to 4 years of age. Age of acquisition ranged from 2 to 6 years of age. Thus, while frequently delayed, the ability to fast-map the meaning of novel words was appropriate to the children's language level.

As children develop, acquisition of new words occurs with increasing frequency through inferring the meaning from a reading context. In one study, both oral and SC deaf students were able to infer an initial meaning for novel words presented in written passages, although better readers were able to fast-map more words than poorer readers (De Villiers and Pomerantz, 1992). This suggests that if children were able to understand the semantic, contextual cues given in these passages, they were able to infer the meaning of the words. On the other hand, acquisition of the grammatical nature of the new words did not occur during the fast-mapping process. Thus, although the children could define the new words, they could not categorize the words as nouns, verbs, or adjectives; nor could they make judgments about the words' correct syntactic usages. Consistent with this finding, in an error analysis of spoken and written English samples, McAfee *et al.* (1990) found that deaf college students used mature vocabulary that was semantically correct but whose form class was incorrect (e.g., "*He tried to suicide himself*," McAfee *et al.*, 1990: 633) In addition, when attempting to describe something they did not have a word for, rather than create complex sentences using their existing lexicon, these students frequently created new word forms that showed considerable semantic creativity but again violated grammatical formation rules (e.g., "*the hang of the neck*, *a hang*, *or hangneck* for the word noose," McAfee *et al.*, 1990: 633)

Learning new words also depends on children's ability to represent or encode the phonological structure of words, and the latter may be a source of slower lexical growth. Gilbertson and Kamhi (1995) found that vocabulary knowledge, as assessed by

the PPVT, related to hard-of-hearing elementary school children's ability to quickly learn to accurately reproduce novel phonologically complex words (e.g., *shabadiffidy*). Hearing children and those hard-of-hearing children who had age-appropriate vocabulary knowledge were able to accurately produce the novel words after only four exposures. On the other hand, half of the hard-of-hearing children had much more difficulty learning, sometimes not producing the complex words correctly even after ten exposures. These latter children performed well below age norms on standardized tests of both receptive and expressive vocabulary.

Thus, slower lexical growth may occur for both exogenous and endogenous reasons. Language impoverished environments result in fewer opportunities to learn new words. In addition, while the ability to fast-map the meaning of words is resilient, deaf children's lexical growth may be hindered by difficulties with quickly learning the phonological representation of words and their syntactic properties.

Content of Lexicon

Even though deaf preschoolers are typically older than hearing children when they are acquiring their initial vocabulary, the words contained in their early lexicons (<35 words) are remarkably similar (Lederberg, 2003). Nominals or names of objects and people predominate in the 50 and 100 word lexicons of deaf and hearing children. However, nominals are relatively less important, with deaf children's lexicons containing more action, descriptive, and personal-social words than those of hearing children at the same level of lexical development.

Older deaf children also showed a delayed and different developmental sequence to their acquisition for particular lexical items (Moeller *et al.*, 1986). On the Boehm Basic Concepts Test, deaf children (6- to 18-year-olds), on average, performed below scores typical of hearing 6-year-olds. In addition, the order of difficulty of the concepts was not the same for deaf and hearing children. For the deaf children, concepts that refer to spatial, size, and perceptual attributes were relatively easy, while time concepts and pronouns were very difficult. In addition, consistent with syntactic deficits, language samples of children with and without cochlear implants are more likely to contain content words and less likely to contain function words compared to hearing children (Schauwers *et al.*, 2005).

Syntactic Development

Deaf children's syntactic development is even more delayed than their vocabulary development. For example, in a study of deaf children ranging in age from 8 to 18 years, age-equivalent scores on receptive and expressive syntax tests averaged from 4 to 6 years old according to hearing norms. In addition, scores did not increase with age (Osberger *et al.*, 1986). The nature of syntactic development is difficult to describe because surprisingly little research has focused on developmental change in deaf children's syntactic knowledge, and given the large individual differences in syntactic performance, cross-study comparisons are difficult. Instead, researchers have focused on characterizing syntactic knowledge of deaf students.

While much of the research has been atheoretical and descriptive in nature, two research reviews (Berent, 1996; De Villiers *et al.*, 1993) suggest deaf individuals' syntactic knowledge can be understood using generative grammar. Similar to hearing children, deaf children acquire lexical categories – noun, verb, adjective, prepositions (or as referred to in other linguistic accounts, thematic categories such as agent, action, patient) – before functional categories. Functional categories are the small class of morphemes that relate thematic categories with one another, specifically, the inflectional, complement, and determiner categories. Even among deaf adolescents and adults, knowledge of functional categories may be fragile or incomplete. Both reviews suggest such incomplete knowledge results in errors in morphosyntactic constructions and complex sentences.

Researchers have confirmed that deaf children acquiring oral or signed English, spoken Hebrew, Italian, and German have particular difficulty with functional aspects of language, including derivational and inflectional markers, and the determiner and auxiliary systems. Children acquiring English (including those in oral and SC environments, with and without cochlear implants; ranging in age from 5 to 20 years) frequently omit tense inflections (present, past, or present progressive) in obligatory contexts (Berent, 1996; De Villiers and Pomerantz, 1992; McAfee *et al.*, 1990; Schauwers *et al.*, in press; Schick, 2003). For instance, in several studies, oral children (6- to 13-year-olds, some with cochlear implants) omitted tense in up to two-thirds of obligatory contexts. Unlike hearing children, deaf children also rarely overgeneralized the regular past tense (-ed) to irregular verbs. Thus, rather than acquiring tense rules, deaf children may have a prolonged period of memorization of verb forms. Another common error for English-speaking and signing children is deletion of the copula in obligatory contexts and failure to supply the correct inflection when present. On the other hand, English deaf children make fewer errors with English plurals and possessives, and unlike tense errors, performance is age-related. Thus, order of

acquisition of these morphemes differs from that by hearing children.

Although all deaf children show morphosyntactic difficulties, which specific morphemes are problematic seem language-dependent. Israeli oral deaf children and German oral deaf children with cochlear implants make few errors on verb tense or noun pluralization (Szagun, 2004; Tur-Kaspa and Dromi, 1998). In fact, German children with cochlear implants performed nearly as well as hearing children in acquiring plurals and verb inflectional morphology (Szagun, 2004). On the other hand, these children made many errors in grammatical agreement. Hebrew children made errors in number and gender agreement between verbs and nouns and between adjectives and nouns (Tur-Kaspa and Dromi, 1998). For example, for the noun phrase, *tmuna levana axat* ('one white picture all feminine'), a student said *tmuna lavan exad* (using the masculine form of 'white' and 'one'). German oral children with cochlear implants also had difficulties with case and gender agreement between articles and nouns (Szagun, 2004). Italian 11- to 15-year-olds made as many errors of agreement between articles and nouns as hearing first graders (Taeschner *et al.*, 1988). Rather than marking agreement for gender and case, these adolescents frequently selected an article based on the final vowel or letter of the head noun, something never seen in hearing children (e.g., *i notti*, instead of *le notti*, *le fucile*, instead of *il fucile*). American deaf children also show problems correctly using discourse agreement rules for determiners and pronouns (De Villiers *et al.*, 1993).

Deaf children also struggle with the auxiliary system (Berent, 1996; De Villiers *et al.*, 1993). Deaf adolescents judge sentences with missing auxiliaries as grammatical. In one study, oral deaf 6- to 14-year-olds only supplied auxiliaries 30% of the time when formulating questions. Unlike hearing children, they almost never produced questions with the auxiliary in the wrong place (e.g., *Where you are going*?), suggesting that the underlying grammar is different for deaf and hearing children. On the other hand, unlike other areas of grammar, deaf college students seemed to have acquired the auxiliary system and rarely made auxiliary errors (McAfee *et al.*, 1990). Thus, while severely delayed, the auxiliary system may eventually be acquired.

Comprehension and production of complex sentences is also problematic (Berent, 1996; De Villiers *et al.*, 1993). Complement structures are frequently misunderstood and not correctly produced even by deaf adolescents. Deaf children and adolescents frequently interpret passive and relative clauses as if they adhere to subject verb object word order. Medial relative clauses are more difficult than final relative clauses and there is little improvement in performance from 6 to 18 years of age. Twenty-five percent of relative clauses produced by oral deaf children were categorized as 'deviant' (i.e., a form hearing children would never produce) (De Villiers *et al.*, 1993). These included redundancy of the resumptive pronoun (e.g., *The girl caught the soccer ball which the boy kick the ball.*), mistakes in relative pronouns (*One the baby is feeding the bird is the mother is carrying on him.*), and modification of the wrong noun phrase (*The boy empty pail who was washing the car*).

Why are these aspects of syntax so difficult for deaf individuals to acquire? Acquisition of syntax seems to depend on input from 'face-to-face' interactions, since deaf adolescents still struggle with these syntactic rules even after years of exposure to them in written materials and formal instruction. This is the case even for children who are fluent in a natural sign language and are learning English as a second language (Mayberry, 2000). In face-to-face interactions, the grammatical elements that are necessary to learn functional categories are unstressed and carry minimal semantic information (De Villiers *et al.*, 1993). Markers such as inflectional morphemes, determiners, and pronouns are less perceptually salient in the speech stream than content words. In addition to accounting for the earlier acquisition of lexical categories, perceptual salience is hypothesized to determine the relative difficulty of different morphemes (e.g., the plural *-s* is perceptually more prominent than the past tense *-ed* and is acquired earlier) (Schauwers *et al.*, 2005). Although intuitively appealing, the little developmental research that exists suggests a more complicated picture, with performance only being imperfectly predicted by perceptual salience (e.g., children acquire plural *-s* but not third person *-s*). Szagun (2004) suggests that syntactic difficulties also can arise from processing deficits in working memory. Because of auditory deprivation during early development, deaf children may find it particularly difficult to acquire elements that require processing and storing of information across elements in the sentence, such as gender and case agreements. Indeed, preliminary research suggests that implantation before age 2 is more likely to result in near normal rates of language development (including syntax) than those after 2 (Nicholas and Geers, 2005; Schauwers *et al.*, 2005). These hypotheses can be evaluated in future research through cross-linguistic analysis of relations between perceptual salience, age of implantation, and syntactic development.

Although MCE was created to make English syntax visible and thus to circumvent the perceptual problems of spoken language, it has clearly not eliminated deaf children's syntactic problems. This may be

partially because less skilled signers who are providing language models for deaf children frequently delete functional morphemes. However, even when children learn language from skilled signers, they still struggle with auxiliaries, copula, and inflectional morphology. Perceptual processes may also interfere with acquiring spoken language through a visual modality (Mayberry, 2000; Schick, 2003). Many of the invented signs that mark functional elements are difficult to visually discriminate because they require the receiver to make fine distinctions using peripheral vision. In addition, visual working memory may not be well suited to learning relations between elements when they are marked sequentially as is the case in MCE, rather than simultaneously as is done by natural sign languages (Spencer, 2004).

Summary and Conclusion

Deaf and hearing toddlers are remarkably similar in the development of joint attention, intentional communication, and the use of nonlinguistic gestures and vocalizations to communicate with others around them. However, differences in language skills are evident even during the first year, when deaf toddlers fail to produce canonical babbling and do not acquire the intonational contours of their ambient language. Although there are wide individual differences, deaf children and adolescents frequently have small lexicons and poor syntactic knowledge, especially in the areas of morphosyntax and complex sentences. Some of these deficits come from restricted input due to hearing loss and, for those in a SC environment, because others fail to use sign fluently. However, even when input becomes accessible, there may be permanent and irreversible effects from language deprivation during the early years of life. Research on early identification and intervention suggests access to language during infancy may result in qualitatively better language outcomes than after infancy. In addition, cochlear implantation before 24 months may result in better reception and expression of speech. New interventions based on knowledge of the visual needs of deaf children may also lead to better SC intervention efforts (Meadow-Orlans *et al*., 2004). Thus, the next generation of deaf children with hearing parents may not exhibit such profound linguistic deficits. However, research on the effects of early intervention and early cochlear implantation is still emerging, so we do not know if the promise of these new developments will become a reality.

See also: Language Development: Overview; Language Education of the Deaf; Sign Language: Acquisition.

Bibliography

Berent G P (1996). 'The acquisition of English syntax by deaf learners.' In Ritchie W C & Bhatia T K (eds.) *Handbook of second language acquisition.* San Diego: Academic Press. 469–506.

Crosson J & Geers A E (2001). 'Analysis of narrative ability in children with cochlear implants.' *Ear & Hearing 22,* 381–394.

De Villiers J, De Villiers P & Hoban E (1993). 'The central problem of functional categories in the English syntax of oral deaf children.' In Tager-Flusberg H (ed.) *Constraints on language acquisition: studies of atypical children.* Hillsdale, NJ: Lawrence Erlbaum Assoc. 8–47.

De Villiers P & Pomerantz S B (1992). 'Hearing-impaired students learning new words from written context.' *Applied Psycholinguistics 13,* 409–431.

Gilbertson M & Kamhi A G (1995). 'Novel word learning in children with hearing impairment.' *Journal of Speech and Hearing Research 38,* 630–642.

Jeanes R C, Nienhuys T G W N & Rickards F W (2000). 'The pragmatic skills of profoundly deaf children.' *Journal of Deaf Studies and Deaf Education 5,* 237–247.

Lederberg A R (2003). 'Expressing meaning: from communicative intent to building a lexicon.' In Marschark M & Spencer P E (eds.) *Oxford handbook of deaf studies, language, and education.* New York: Oxford University Press. 247–260.

Lederberg A R & Everhart V S (2000). 'Conversations between deaf children and their hearing mothers: pragmatic and dialogic characteristics.' *Journal of Deaf Studies and Deaf Education 5(4),* 303–322.

Lederberg A R, Prezbindowski A K & Spencer P E (2000). 'Word learning skills of deaf preschoolers: the development of novel mapping and rapid word learning strategies.' *Child Development 53,* 1055–1065.

Lederberg A R & Spencer P E (2005). 'Critical periods in the acquisition of lexical skills: evidence from deaf individuals.' In Fletcher P & Miller J F (eds.) *Language disorders and developmental theory.* Philadelphia, PA: John Benjamins. 121–145.

Marschark M (2001). *Language development in children who are deaf: A research synthesis.* Alexandria, VA: National Association of State Directors of Special Education.

Marschark M, Mouradian V & Halas M (1994). 'Discourse rules in the language production of deaf and hearing children.' *Journal of Experimental Child Psychology 57,* 89–107.

Mayberry R I (2004). 'Cognitive development in deaf children: the interface of language and perception in cognitive neuroscience.' In Rapin & Segalowitz S (eds.) *Child neuropsychology, vol. 7 of the Handbook of neuropsychology.* Boller F & Grafman J (series eds.), 2nd edn. Amsterdam: Elsevier. 71–107.

Mayne A M, Yoshinaga-Itano C, Sedey A L & Carey A (2000). 'Expressive vocabulary development of infants and toddlers who are deaf or hard of hearing.' *The Volta Review 100(5),* 1–28.

McAfee M, Kelly J F & Samar V J (1990). 'Spoken and written English errors of postsecondary students with

severe hearing impairment.' *Journal of Speech and Hearing Disorders 55,* 628–634.
Meadow-Orlans K P, Spencer P E & Koester L S (2004). *The world of deaf infants: a longitudinal study.* New York: Oxford University Press.
Moeller M P (2000). 'Early intervention and language development in children who are deaf and hard of hearing.' *Pediatrics 106(3),* e43 (41–49).
Moeller M P, Osberger M J & Eccarius M (1986). 'Receptive language skills.' *Language and Learning Skills in Hearing-Impaired Children 23,* 41–53.
Nicholas J G (2000). 'Age differences in the use of informative/heuristic communicative functions in young children with and without hearing loss who are learning spoken language.' *Journal of Speech, Language, and Hearing Research 43,* 380–394.
Nicholas J G & Geers A E (2005). 'The process and early outcomes of cochlear implantation by three years of age.' In Spencer P E & Marschark M (eds.) *Advances in the spoken language development of deaf children.* New York: Oxford University Press.
Oller D K & Eilers R E (1988). 'The role of audition in infant babbling.' *Child Development 59(2),* 441–449.
Osberger M J, Moeller M P, Eccarius M, Robins A M & Johnson D (1986). 'Expressive language skills.' *Monographs of the American Speech, Language, and Hearing Association 23,* 54–65.
Prezbindowski A K, Adamson L B & Lederberg A R (1998). 'Joint attention in deaf and hearing 22-month-old children and their hearing mothers.' *Journal of Applied Developmental Psychology 19(3),* 377–387.
Schauwers K, Gillis S & Govaerts P (2005). 'Language acquisition in children with a cochlear implant.' In Fletcher P & Miller J F (eds.), *Language disorders and developmental theory.* Philadelphia, PA: John Benjamins. 95–119.
Schick B (2003). 'The development of American Sign Language and manually coded English systems.' In Marschark M & Spencer P E (eds.) *Oxford handbook of deaf studies, language, and education.* New York: Oxford University Press. 219–231.
Spencer P E (1993). 'Communication behaviors of infants with hearing loss and their hearing mothers.' *Journal of Speech and Hearing Research 36,* 311–321.
Spencer P E (2004). 'Individual differences in language performance after cochlear implantation at one to three years of age: child, family, and linguistic factors.' *Journal of Deaf Studies and Deaf Education 9,* 393–412.
Stoel-Gammon C (1988). 'Prelinguistic vocalizations of hearing-impaired and normally hearing subjects: a comparison of consonantal inventories.' *Journal of Speech and Hearing Disorders 53,* 302–315.
Szagun G (2004). 'Learning by ear: on the acquisition of case and gender marking by German-speaking children with normal hearing and with cochlear implants.' *Journal of Child Language 31,* 1–30.
Taeschner T, Devescovi A & Volterra V (1988). 'Affixes and function words in the written language of deaf children.' *Applied Psycholinguistics 9,* 385–401.
Tait M, Lutman M E & Robinson K (2000). 'Preimplant measures of preverbal communicative behavior as predictors of cochlear implant outcomes in children.' *Ear and Hearing 21,* 18–24.
Tur-Kaspa H & Dromi E (1998). 'Spoken and written language assessment of orally trained children with hearing loss: syntactic structures and deviations.' *Volta Review 100,* 186–202.
Wallace V, Menn L & Yoshinaga-Itano C (2000). 'Is babble the gateway to speech for all children? A longitudinal study of children who are deaf or hard of hearing.' *The Volta Review 100(5),* 121–148.
Yoshinaga-Itano C, Sedey A L, Coulter D K & Mehl A L (1998). 'Language of early- and late-identified children with hearing loss.' *American Academy of Pediatrics 102(5),* 1161–1171.

Language Development in School-Age Children, Adolescents, and Adults

M A Nippold, University of Oregon, Eugene, OR, USA

In this article, language development beyond the preschool years is described, covering significant changes that occur in school-age children (ages 6–12 years), adolescents (ages 13–19 years), and adults (ages 20+ years). For many years, it was widely believed that language development was largely complete by the time a child entered school, at five or six years of age. More recently, however, it has become apparent that a great deal of linguistic growth occurs during the school-age years, adolescence, and well into adulthood, and that the development of language has no obvious point of completion (Nippold, 1998). This should not detract from the undeniable truth that language development in infants, toddlers, and preschool children is an amazing phenomenon, marked by rapid and salient change. By contrast, language growth beyond the preschool years is comparatively slow, subtle, and sometimes difficult to quantify. Nevertheless, later language development is an exciting area of scientific inquiry.

The differences between early and later language development are striking. For example, after steadily

Nippold M A, Leonard L B & Kail R (1984). 'Syntactic and conceptual factors in children's understanding of metaphors.' *Journal of Speech and Hearing Research 27,* 197–205.
Nippold M A & Martin S T (1989). 'Idiom interpretation in isolation versus context: a developmental study with adolescents.' *Journal of Speech and Hearing Research 32, 59–66.*
Nippold M A, Martin S A & Erskine B J (1988). 'Proverb comprehension in context: a developmental study with children and adolescents.' *Journal of Speech and Hearing Research 31,* 19–28.
Nippold M A, Moran C & Schwarz I E (2001). 'Idiom understanding in pre-adolescents: synergy in action.' *American Journal of Speech–Language Pathology 10,* 169–179.
Nippold M A & Rudzinski M (1993). 'Familiarity and transparency in idiom explanation: a developmental study of children and adolescents.' *Journal of Speech and Hearing Research 36,* 728–737.
Nippold M A, Taylor C L & Baker J M (1996). 'Idiom understanding in Australian youth: a cross-cultural comparison.' *Journal of Speech and Hearing Research 39,* 442–447.
Nippold M A, Ward-Lonergan J M & Fanning J L (2005b). 'Persuasive writing in children, adolescents, and adults: a study of syntactic, semantic, and pragmatic development.' *Language, Speech, and Hearing Services in Schools 36,* 125–138.
Paul R (1981). 'Analyzing complex sentence development.' In Miller J F (ed.) *Assessing language production in children: experimental procedures.* Baltimore, MD: University Park Press. 36–40.
Raffaelli M & Duckett E (1989). '"We were just talking . . .": conversations in early adolescence.' *Journal of Youth and Adolescence 18, 567–582.*
Rawlins W K (1992). *Friendship matters: communication, dialectics, and the life course.* NY: DeGruyter.
Scott C M (1988). 'Spoken and written syntax.' In Nippold M A (ed.) *Later language development: ages nine through nineteen.* Boston: College-Hill. 49–95.
Stanovich K E & Cunningham A E (1992). 'Studying the consequences of literacy within a literate society: the cognitive correlates of print exposure.' *Memory and Cognition 20(1),* 51–68.
Sternberg R J (1987). 'Most vocabulary is learned from context.' In McKeown M G & Curtis M E (eds.) *The nature of vocabulary acquisition.* Hillsdale, NJ: Erlbaum. 89–105.
Verhoeven L, Aparici M, Cahana-Amitay D, van Hell J, Kriz S & Viguie-Simon A (2002). 'Clause packaging in writing and speech: a cross-linguistic developmental analysis.' *Written Language and Literacy 5(2),* 135–162.

Language Development: Morphology

H Behrens, Rijksuniversiteit, Groningen, The Netherlands

Most studies on the acquisition of morphology focus on inflectional morphology (but see Clark, 1993 for results on compounding and derivation). Of particular interest are studies on the typical order of acquisition of morphology, as well as on individual and crosslinguistic differences and the causes underlying such variation.

In many languages, the earliest stages of language development show characteristics of 'telegraphic speech.' That is, inflectional morphemes and function words are initially absent but are gradually added on in the process of acquisition.

Studies of Morpheme Order Acquisition

Brown (1973: 271, Figure 14) provided a thorough methodological milestone for the systematic comparison of individual developmental trajectories. Brown related the acquisition of 14 carefully selected grammatical morphemes to stages of development, as defined by the mean length of utterance (MLU), and productivity, as defined by 90% of use in obligatory contexts. Upon a systematic scoring, the three English-speaking children under investigation showed a high correlation in their acquisition of the 14 morphemes. Present progressive was typically the first morpheme to become productive, followed by the prepositions *in/on*, plural and past irregular marking (Brown, 1973: 274, Table 38). Brown's results corresponded to the results from previous diary studies and experimental data (Brown, 1973: 275ff.) and inspired a wide range of comparative research.

Individual Differences

The availability of more and more data on individual developmental trajectories revealed not only that children differ in the age at which they acquire particular morphemes, but also that children show qualitatively different styles of language acquisition regarding the units they first pick up from the input sound stream. While most children initially identify stressed stems without unstressed inflectional elements, other

children show a tendency to use larger, unanalyzed chunks and are faithful to the prosodic contour instead. This distinction is known under dichotomies such as "referential vs. expressive" or "analytic vs. holistic" (Peters, 1995; Lieven, 1997).

However, crosslinguistic research pointed out that holistic vs. analytic processing is not merely a matter of individual style, but also much influenced by language type.

Crosslingustic Differences

English often serves as a benchmark for acquisition processes, although it is a rather atypical example for the acquisition of morphology. English has little inflectional morphology, and in most cases the morphological word consists of just the uninflected stem. In the past three decades, crosslinguistic research has demonstrated that the course of morphological acquisition critically relies on the language type. English represents an almost isolating language in which acquisition is a matter of additive complexity. The child can easily identify the stem and then acquire inflectional morphemes as additional material. In other languages with richer and more complex inflectional morphology, children first have to identify the stem and the inflectional processes that act on the stem. Thus, in many languages children first use inflected forms in an unanalyzed, formulaic way and then have to break up and reanalyze the complex form in order to build morphological paradigms (see Voeikova and Dressler, 2002 for comparative studies on paradigm building).

The acquisition of morphology is related to language type. Slobin (1973) demonstrated that agglutinative languages like Turkish with very regular form-function matches provide the fewest problems, as opposed to more irregular inflectional languages like Serbo-Croatian. In incorporating or polysynthetic languages, children virtually never start out with uninflected lexical roots. The degree to which grammatical morphemes are easily identifiable influences not only which acquisition style is most successful (Lieven, 1997), but also the age at which morphology is acquired. The more regular and semantically transparent the inflection, the earlier its acquisition. For detailed studies on the acquisition of particular morphemes in particular languages, consult the online CHILDES bibliography. Thorough surveys on a set of languages are presented in Slobin, 1985a, 1992, and 1997a.

Morphology as an Interface

Current handbooks and textbooks on language acquisition often lack a chapter on the acquisition of morphology. Studies typically focus on the syntactic function of morphology, the semantics of morphemes, or the nature of psycholinguistic processes. Consequently, the acquisition of morphology is treated in chapters on syntactic or semantic development or in sections on processing.

Morphology and Syntax

Within generative linguistics, the acquisition of morphology as such does not figure prominently. Instead, morphological categories are relevant in their function of licensing syntactic processes, for example, the projections of functional categories in the principles and parameters framework (e.g., Radford, 1995) or the merge operations associated with strong and weak features in minimalism. However, there is some debate regarding the architecture of the human language processing system concerning the question of whether words and formal structures (rules) are processed by different mechanisms (see below).

Morphology and Semantics

Space and time are well-studied semantic notions crosslinguistically. For spatial morphology (free and bound) language specific trajectories are proposed. For example, children learning Korean categorize spatial relations differently from children acquiring English, who again differ from children learning Dutch. That is, there are no universally privileged notions, and the acquisition of spatial language is now one of the strongholds for linguistic relativity, which claims that language structure influences cognitive categorization (Bowerman, 1996). For temporal reference, in contrast, there is a strong crosslinguistic tendency that early past tense marking is associated with resultative verbs (Li and Shirai, 2000). This preference is due to the distributional bias found in most input languages. Across several semantic domains, there is now accumulating evidence that children acquire language-specific notions early, without starting out with a set of possibly universal semantic distinctions first (Slobin, 1997b).

Representation and Processing

Measures of Productivity

Ever since the first systematic studies of child language, the difficulty of assessing children's underlying linguistic competence has been a methodological problem, because not everything children produce is productive, and because reaching full linguistic productivity is a gradual process. Although children's productions may appear to be correct, the use of a complex form could be unanalyzed and formulaic, or

the form might not yet encode the full semantic range of applications.

In spontaneous data, overgeneralization errors are the clearest sign of productivity. Here, children create their own regular form instead of using the irregular form (e.g., *go-ed* instead of *went*). But there are other ways of assessing (degrees of) productivity. In corpus studies, several measures can be used to define (a) first productivity and (b) degrees of mastery. The onset of productivity defines the moment when a given item is no longer used in a lexically specific, frozen, and formulaic way. Morphological contrast is used to demonstrate that the child has analyzed the complex nature of a word form. The child either starts to use a particular morpheme with *x* number of different stems or uses a particular stem with *x* different grammatical morphemes such that morphological paradigms are emerging. Full productivity is attested when the child uses the structure in the majority of appropriate contexts. Commonly, a criterion higher than 75% is used.

Experiments provide a more clear-cut way of assessing the child's morphological competence under controlled conditions. Tests that use nonce-word material are one successful measure of productivity. In 1958, Jean Berko published the results of the so-called wug test. *Wug* was one of the names of fantasy animals and actions. Children had to learn these nonce words and then had to pluralize them (e.g., *wug-s*) or to form the past tense (e.g., *loodge-d*). The use of new lexical material guarantees that the child cannot retrieve a holistically stored form from memory, but instead has to inflect the new word spontaneously.

The range of productivity measures indicates that different linguistic domains or research interests require different thresholds. The reader should pay attention to the productivity criteria used when comparing the results of different studies. A survey of the average age of productivity of different morphological and syntactic structures is presented in Tomasello (2000).

Processing

From a psycholinguistic perspective, it is of interest how morphological paradigms and inflected forms are stored and processed. There is substantial work on the acquisition of regular and irregular allomorphs. The debate between the dual- and the single-mechanism model has received much attention since the mid-1980s (cf. Clahsen, 1999 and commentaries). It centers on the question of whether regular and irregular morphology is stored separately or not.

Representational approaches The dual-route model is based the idea that morphology is processed by different mechanisms and predicts that irregular forms are stored in memory in a holistic fashion, whereas regular morphology is processed by a symbolic rule such that stem and affixes are stored analytically. Overgeneralization errors are the result of memory retrieval failure: In the encoding process, the child first looks up irregular forms in the mental lexicon, but if this fails, regular morphology is applied instead.

Processing models In contrast, single-mechanism models of various kinds rely on general processing mechanisms that achieve the acquisition task without requiring specifically linguistic representations or symbolic rules. The single-mechanism model for inflection originated in connectionism (Rumelhart and McClelland, 1987; Plunkett, 1995), but similar ideas are prominent also within noncomputational models. First, Slobin (1985b) formulated a set of "operating principles," some of which were segmentation strategies to identify morphemes (e.g., "pay attention to the end of words"). Later approaches clarified the ranking and the nature of interaction of such processing strategies. The competition model (Bates and MacWhinney, 1989) is designed to handle multifactorial grammatical phenomena because it is assumed that these systems offer multiple cues (e.g., inflection plus word order) that differ in strength (perceptual salience, frequency) and validity (non-ambiguity). The stronger and the more reliable the cue, the easier the acquisition of such a cue. Research with infants has demonstrated that they are able to identify stochastic patterns even in the prelinguistic period (Aslin *et al.*, 1999).

See also: Associationism and Connectionism; CHILDES Database; Computational Approaches to Language Acquisition; Productivity; Register Variation: Core Grammar and Periphery; Syntactic Development.

Bibliography

Aslin R N, Saffran J R & Newport E L (1999). 'Statistical learning in linguistic and nonlinguistic domains.' In MacWhinney B (ed.) *The emergence of language.* Mahwah, NJ: Erlbaum. 359–380.

Bates E & MacWhinney B (1987). 'Competition, variation and language learning.' In MacWhinney B (ed.) *Mechanisms of language acquisition.* Hillsdale, NJ: Erlbaum. 157–193.

Berko J (1958). The child's learning of English morphology. *Word 14,* 150–177.

Bowerman M (1996). 'Learning how to structure space for language: a crosslinguistic perspective.' In Bloom P, Peterson M A, Nadel L & Garrett M F (eds.) *Language and space*. Cambridge, MA: MIT Press. 385–436.
Brown R (1973). *A first language: the early stages*. Cambridge, MA: Harvard University Press.
Clahsen H (1999). Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioral and Brain Sciences 22,* 991–1060.
Clark E V (1993). *The lexicon in acquisition*. Cambridge: Cambridge University Press.
Li P & Shirai Y (2000). *The acquisition of lexical and grammatical aspect*. Berlin: de Gruyter.
Lieven E V M (1997). 'Variation in a crosslinguistic context.' In Slobin D I (ed.) *The crosslinguistic study of language acquisition. Vol. 5: Expanding the contexts*. Mahwah, NJ: Erlbaum. 199–263.
Peters A M (1995). 'Strategies in the acquisition of syntax.' In Fletcher P & MacWhinney B (eds.) *The handbook of child language*. Oxford: Blackwell. 462–482.
Plunkett K (1995). 'Connectionist approaches to language acquisition.' In Fletcher P & MacWhinney B (eds.) *The handbook of child language*. Oxford: Blackwell. 36–72.
Radford A (1995). 'Phrase structure and functional categories.' In Fletcher P & MacWhinney B (eds.) *The handbook of child language*. Oxford: Blackwell. 483–507.
Rumelhart D E & McClelland J L (1987). 'Learning the past tense of English verbs: Implicit rules or parallel processing?' In MacWhinney B (ed.) *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum. 195–248.
Slobin D I (1973). 'Cognitive prerequisites for the development of grammar.' In Ferguson C A & Slobin D I (eds.) *Studies of child language development*. New York: Holt, Rinehart & Winston. 175–208.
Slobin D I (ed.) (1985a). *The crosslinguistic study of language acquisition. Vol. 1: The data*. Hillsdale, NJ: Erlbaum.
Slobin D I (1985b). 'Crosslinguistic evidence for the language-making capacity.' In Slobin D I (ed.) *The crosslinguistic study of language acquisition. Vol. 2: Theoretical issues*. Hillsdale, NJ: Erlbaum. 1157–1249.
Slobin D I (ed.) (1992). *The crosslinguistic study of language acquisition. Vol. 3*. Hillsdale, NJ: Erlbaum.
Slobin D I (ed.) (1997a). *The crosslinguistic study of language acquisition. Vol. 4*. Mahwah, NJ: Erlbaum.
Slobin D I (1997b). The origins of grammaticizable functions: beyond the individual mind. In Slobin D I (ed.) *The crosslinguistic study of language acquisition. Vol. 5: Expanding the contexts*. Mahwah, NJ: Erlbaum. 265–323.
Tomasello M (2000). Do your children have adult syntactic competence? *Cognition 74,* 209–253.
Voeikova M D & Dressler W U (eds.) (2002). *Pre- and protomorphology: early phases of morphological development in nouns and verbs*. München: Lincom Europa.

Relevant Website

http://childes.psy.cmu.edu – CHILDES (Child Language Data Exchange System) bibliography.

Language Development: Overview

E Lieven, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

All over the world, children learn to talk on a roughly equivalent timetable. They do so by learning the language or languages of their environment. There is considerable debate over what cognitive, social, or specifically linguistic, innate capacities they bring to language leaning. This article begins with a brief timetable of development and then focuses in turn on the major aspects of language learning in terms of infancy, learning words, learning morphology, early grammar, later grammar, and the learning of pragmatic and metalinguistic skills. It concludes with some brief reflections on atypical development. The relevant theoretical issues are covered as they arise in each section and are considered again in the last section on learnability and constituency.

During infancy, children develop a wide range of cognitive and social skills together with a developing ability to segment the speechstream into meaningful units. They usually produce their first recognizable words somewhere between 10–18 months of age and their first multiword utterances between 14–24 months. By age 3, children are often able to produce quite long utterances and are beginning to be able to combine more than one clause into coordinate and subordinate constructions (e.g., relative clauses, cleft sentences). Between the ages of 4 to 7, there are major advances in children's ability to take the perspective of the listener into account and to produce coherent discourse and narrative sequences. These abilities, as well as the ability to reflect on language as an object of knowledge, develop throughout the school years and are much influenced by the extent of literacy or other complex language (for instance, ritual language) to which children are exposed.

Bowerman M (1996). 'Learning how to structure space for language: a crosslinguistic perspective.' In Bloom P, Peterson M A, Nadel L & Garrett M F (eds.) *Language and space*. Cambridge, MA: MIT Press. 385–436.
Brown R (1973). *A first language: the early stages*. Cambridge, MA: Harvard University Press.
Clahsen H (1999). Lexical entries and rules of language: a multidisciplinary study of German inflection. *Behavioral and Brain Sciences 22,* 991–1060.
Clark E V (1993). *The lexicon in acquisition.* Cambridge: Cambridge University Press.
Li P & Shirai Y (2000). *The acquisition of lexical and grammatical aspect.* Berlin: de Gruyter.
Lieven E V M (1997). 'Variation in a crosslinguistic context.' In Slobin D I (ed.) *The crosslinguistic study of language acquisition. Vol. 5: Expanding the contexts.* Mahwah, NJ: Erlbaum. 199–263.
Peters A M (1995). 'Strategies in the acquisition of syntax.' In Fletcher P & MacWhinney B (eds.) *The handbook of child language.* Oxford: Blackwell. 462–482.
Plunkett K (1995). 'Connectionist approaches to language acquisition.' In Fletcher P & MacWhinney B (eds.) *The handbook of child language.* Oxford: Blackwell. 36–72.
Radford A (1995). 'Phrase structure and functional categories.' In Fletcher P & MacWhinney B (eds.) *The handbook of child language.* Oxford: Blackwell. 483–507.
Rumelhart D E & McClelland J L (1987). 'Learning the past tense of English verbs: Implicit rules or parallel processing?' In MacWhinney B (ed.) *Mechanisms of language acquisition.* Hillsdale, NJ: Erlbaum. 195–248.
Slobin D I (1973). 'Cognitive prerequisites for the development of grammar.' In Ferguson C A & Slobin D I (eds.) *Studies of child language development.* New York: Holt, Rinehart & Winston. 175–208.
Slobin D I (ed.) (1985a). *The crosslinguistic study of language acquisition. Vol. 1: The data.* Hillsdale, NJ: Erlbaum.
Slobin D I (1985b). 'Crosslinguistic evidence for the language-making capacity.' In Slobin D I (ed.) *The crosslinguistic study of language acquisition. Vol. 2: Theoretical issues.* Hillsdale, NJ: Erlbaum. 1157–1249.
Slobin D I (ed.) (1992). *The crosslinguistic study of language acquisition. Vol. 3.* Hillsdale, NJ: Erlbaum.
Slobin D I (ed.) (1997a). *The crosslinguistic study of language acquisition. Vol. 4.* Mahwah, NJ: Erlbaum.
Slobin D I (1997b). The origins of grammaticizable functions: beyond the individual mind. In Slobin D I (ed.) *The crosslinguistic study of language acquisition. Vol. 5: Expanding the contexts.* Mahwah, NJ: Erlbaum. 265–323.
Tomasello M (2000). Do your children have adult syntactic competence? *Cognition 74,* 209–253.
Voeikova M D & Dressler W U (eds.) (2002). *Pre- and protomorphology: early phases of morphological development in nouns and verbs.* München: Lincom Europa.

Relevant Website

http://childes.psy.cmu.edu – CHILDES (Child Language Data Exchange System) bibliography.

Language Development: Overview

E Lieven, Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

All over the world, children learn to talk on a roughly equivalent timetable. They do so by learning the language or languages of their environment. There is considerable debate over what cognitive, social, or specifically linguistic, innate capacities they bring to language leaning. This article begins with a brief timetable of development and then focuses in turn on the major aspects of language learning in terms of infancy, learning words, learning morphology, early grammar, later grammar, and the learning of pragmatic and metalinguistic skills. It concludes with some brief reflections on atypical development. The relevant theoretical issues are covered as they arise in each section and are considered again in the last section on learnability and constituency.

During infancy, children develop a wide range of cognitive and social skills together with a developing ability to segment the speechstream into meaningful units. They usually produce their first recognizable words somewhere between 10–18 months of age and their first multiword utterances between 14–24 months. By age 3, children are often able to produce quite long utterances and are beginning to be able to combine more than one clause into coordinate and subordinate constructions (e.g., relative clauses, cleft sentences). Between the ages of 4 to 7, there are major advances in children's ability to take the perspective of the listener into account and to produce coherent discourse and narrative sequences. These abilities, as well as the ability to reflect on language as an object of knowledge, develop throughout the school years and are much influenced by the extent of literacy or other complex language (for instance, ritual language) to which children are exposed.

Throughout this developmental timetable, there are major individual differences in the ages at which children reach these points and, in addition, in the balance of skills that a particular child may manifest at a particular point in time. There are also individual differences in how children tackle any of these tasks. This is an important point to remember when considering theories that rely for their confirmation on a particular order of development or on a particular relationship among different skills. It is also important to remember that many children (perhaps most) grow up hearing and, to some extent at least, learning more than one language. The evidence to date is that doing so does not have a significant impact on the developmental timetable for language learning in the early years.

Overview of Development

Infancy

Children are born with the ability to discriminate their mother's voice from that of other women and to discriminate speech from nonspeech, presumably because of their experience in the womb. In the earliest months of life, they seem able to discriminate sounds produced in the languages of the world, but this ability diminishes until, by 8 to10 months, their ability to discriminate between sounds is confined to the sounds of their native language(s).

Experiments with children between 6 to 10 months of age indicate a developing sensitivity to the prosodic indicators of major phrasal units in the speechstream. Thus, 9-month-old infants prefer to listen to speech that is segmented using pauses at major clause boundaries, rather than within clauses, whereas there is no difference in listening preferences for 6-month-olds. Around 7.5 months infants also become able to identify words in the speechstream. Experiments with English-, Japanese-, and French-learning children have shown that children can discriminate words that they have heard before from those that they have not, even when the words were embedded in speech. They can also discriminate highly frequent words in their language from low-frequency words, and words that are typically learned early by children from words that are not typically learned early (Jucszyk, 1997).

Experiments with children about 10 months of age have indicated that they are sensitive to the ordering of words in their language, thus being able to discriminate normal English sentences from ones in which determiners and nouns were reversed (e.g., *kitten the*) or ones in which some grammatical morphemes were replaced by nonsense syllables. Older children (at 18 months but not at 15 months) can discriminate sentences in which the combination of auxiliary and verb was correct (*is running*) from those in which it was not (*can running*). Finally, using simple artificial languages, it has been claimed that infants between 7.5 to 17 months are able to recognize strings with the same ordering rules, but with different 'vocabulary' after a short exposure. However, it seems that this discrimination is possible only when there are patterns of repetition in the 'vocabulary items' (Gomez *et al*., 2000).

These experiments show a clear path of development during the first 12 to 18 months of life. As children's experience with language develops, so do their segmentation, word recognition, and pattern recognition skills. These skills are obviously central to the child's ability to parse the input and to start to connect it to meaning. It is important to note that all these experiments depend on the infant's ability to discriminate one stimulus from another. They do not have to understand or use the stimuli in communication nor connect them to any meaning in the environment.

However, infants do make huge developmental strides in their cognitive and social development during the first 12 months of life. Cognitively, infants (probably from birth) have clear expectations about the ways in which objects will behave, and these expectations develop in sophistication over the first year of life. Before their first birthday, they are able to form categories of objects based both on form and, to some extent, on function (Mandler, 2000).

During the first 6 to 8 months of life, children develop the ability to interact with others, to make demands, and to resist them. At around 8 to 9 months, there is a major 'step change' as the infant starts to understand that others have intentions that may be different from their own and to incorporate this understanding of other minds into their behavior; for instance, imitating the perceived intentions of others and showing objects that they know their interactant has not seen (Carpenter *et al*., 1998). Infants aged 12 to18 months are clearly starting to associate attention sharing, demanding, and assisting others with specific gestures that they use and, also perhaps, with systematic vocalizations. This development is highly correlated with word learning, as it seems to be the beginning of the attempt to match form to meaning.

Thus, by about age 10 to 12 months, infants are ready to put the patterns skills that they have extracted together with their communicative and meaning-inference skills. It is this combination that is the true start of human language learning.

How many of these developing abilities are unique to humans, and how many are shared with the other

great apes or primates? This is obviously a contentious issue, and researchers of different theoretical persuasions have made very different claims. Much more research remains to be done to tease apart the exact nature of the skills and their precise developmental sequence. However, a number of experiments have indicated that many of the pattern recognition skills shown by human infants may be shared by other primates. In contrast, the intention-reading and communicative skills shown by infants in the last trimester of their first year do seem uniquely human, though nonhuman primates show some precursors of these abilities, especially those few who have been reared in a 'human' environment (see Tomasello, 2003).

Early Comprehension and Production

The time when children first start to show signs of comprehension and to start producing words varies. However, because much depends on how comprehension and production are defined, it is difficult to give an exact range of variation. Comprehension of individual words such as 'No' starts very early, and children can use contextual cues to interpret the utterances of those around them without parsing much of what they hear. Although most well-controlled studies of early word learning have found that comprehension is achieved in advance of production and starts early, production is more variable, with some children producing their first words (as reported by parents) at around 10 months, whereas others might not produce more than a few recognizable words before 16 to 18 months (Fenson *et al.*, 1994). For most children, progress in comprehension and in production is highly correlated, but there are reports of some children whose comprehension outstrips their production by much more than the normal extent (Bates *et al.*, 1988).

Studies have shown that, even when infants know a word, it will take 15-month-olds much longer to process it and to look to a matching picture than it will older infants. Thus, 15-month-olds need to hear the whole word before looking to the matching picture, whereas 18- and 24-month-olds process it 150 and 300 msec faster, respectively. These two older groups shift their gaze before the word has ended, using the ongoing phonological information to discriminate the matching picture from a distractor (Fernald *et al.*, 1998). This development of automatization is clearly one of the major processes taking place between the earliest word-meaning mappings and the point where vocabulary learning develops rapidly.

The production of utterances also goes through a long process of automatization and refinement. Initially, infants' utterances can match the adult form quite closely. However, they go through a subsequent stage in which each child produces a wider range of words, most of which are reduced to a phonological 'template.' The nature of this template varies from child to child, though it always bears a phonological and segmental relationship to the particular language being learned. Children vary greatly in how accurately they can produce the phonology of their language and in how long it takes for this accuracy to develop. They may also differ early on in the extent to which they pick up on the major tunes of the language, as some children tend to produce shorter and more well-articulated utterances (for detailed summaries, see Lieven, 1997; Peters, 1997).

Some early words are highly context dependent – for example, Piaget's classic example of his daughter, Jacqueline who used *voua-ou* (wow-wow) to refer to everything that could be seen from the balcony – but it seems to be the case that children also produce relatively context-independent words from early on (e.g., *more* to request a number of items, not just food). Context dependence is similar to 'underextension' in which the child uses the word to refer to a limited set of referents. This contrasts with overextension where the word is used to refer to a wider group of referents than in the adult language (an example is the use of *Daddy* to refer to all men whom the child encounters). Much early word comprehension and production shows both characteristics and initially is not stable (see Clark, 2003).

One of the most striking aspects of very early speech is its phonological inaccuracy. Analysis makes it clear that the phonological errors that children produce are systematic. Individual children drop syllables or substitute a particular sound of their own for sounds in their language (for instance, *Puggle* for *Puddle*) while still being able to hear the difference between their own production repeated to them and the 'real' word. Children will not, in fact, accept a repetition of their version of a word as being the 'real' word. This phenomenon raises interesting issues for the relationship between comprehension and production. There is major theoretical debate, however, concerning how it should be analyzed. The debate is between a 'templatic approach,' which characterizes the process as the child assimilating words to a production template (Vihman, 1996) or an optimality theoretical approach that sees the child as seeking to resolve a set of constraints on production, such as 'faithfulness' (i.e., getting as close to the model as possible) and 'markedness' (Bernhardt and Stemberger, 1998).

A second issue is the extent to which children's communicative behavior in these early stages is already 'linguistic' in the sense of being conventionalized, context free, and combinatorial. Thus, researchers

have suggested that children's early communicative acts (e.g., showing as a precursor of declaratives and demanding as a precursor of imperatives) develop into 'protolinguistic' symbols; that is, each child uses a specific set of sounds and gestures to convey different meanings (Bates, 1976).

Learning a Vocabulary and Developing Meaning

After an often slow start, children's vocabulary grows exponentially, with new words being added at a rate of 9 to 10 words per day between the ages of 2 and 6. In industrialized and urbanized societies, children add about 3000 words per year for each year in school. Vocabulary size is related closely to the amount of Child-Directed Speech (CDS) in the early years and to various measures of family talk and schooling in the later years (Fenson *et al.*, 1994; Hart and Risley, 1995).

For children learning many languages, including English, nouns form the largest single category of words, but children also learn a wide range of words from other categories (e.g., *want, me, no, what*). The extent to which nouns predominate early on varies among children learning the same language and among children learning different languages. Young children, under certain pragmatic conditions, require only a minimal number of exposures to a novel noun and its referent to learn its meaning (called 'fast mapping'; Carey and Bartlett, 1978; Clark, 2003).

One issue concerns when children form categories of the words that they are learning. The answer to this depends, in part, on how a category is defined, but evidence suggests that a category of nouns develops relatively early in many languages; that English-speaking children can substitute the determiner *a* for *the* early (while taking many years to learn the full range of determiners and their scope); and that fully fledged categories of verb and adjective take more time to develop.

When children learn a new word, what meaning do they give it? It could be the meaning the word has in the language that they are learning or a meaning based on their cognitive categorizations of the world. For instance, children the world over might already have developed nonverbal categories of 'in-ness' and 'on-ness' during infancy, and when they learn words that express the relation between objects, they might initially match these words to these pre-existent categories. There has been considerable research on this issue, and it seems that this is not what happens. Instead, children seem to learn the word and its referent together. Thus, English-learning children use the prepositions *in* to refer to one object contained within another and *on* to refer to one object supported by another, irrespective of how tight the fit is between the objects. However, Korean-learning children use the verb *kkita* to refer to one object inside another only if the fit is tight (e.g., a cassette in a box) and for other tight fit relations between objects (e.g., one object stuck to another with Velcro). The verb *nehta* is used to refer to a loose fit, regardless of whether one object is inside another or on top of it. Children do under- and over-generalize, but they seem to do so along the lines of categorization indicated by their language (Bowerman and Choi, 2001).

How does a child know what someone is referring to when that person uses a word? In principle, the speaker could be referring to any part of the scene. This problem of reference was elucidated most clearly by the philosopher, Quine. The general answer is to suggest that children come to word learning with some already pre-given interpretative skills. One group of researchers has suggested that these skills are in the nature of innate constraints; for instance, the child initially assumes that a word applies to the whole object and that a novel word will not apply to an object for which the child already has a word (mutual exclusivity constraint; Markman *et al.*, 2003). Another group suggests that these skills derive from the pragmatics of the situation that the child already knows about. A series of experiments indicating that children pre-verbally know what is relevant to a situation and to the perspective of another and that, simply through using context, they can find the referent for a word they have never previously seen paired with the referent, suggest that the latter explanation is more likely to be correct (for this debate, see Clark, 2003; Tomasello, 2003).

Another major debate in the literature is over the status of nouns as the first category to be learned. Gentner (1982) has claimed that nouns are universally easier to learn and that they initiate the process of category learning. She argued that this occurs not only because the use of nouns tends to be very frequent in CDS but also because their reference is more transparent – to concrete, imageable objects – and children already have developed a considerable knowledge of objects and their behavior pre-verbally. In opposition to this theory, researchers have argued that children learn a range of words, not just nouns, from the beginning and that there are languages, such as Korean, in which the proportion of nouns in the lexicon does not exceed the proportion of verbs (more or less broadly defined). This issue has not been resolved: on the one hand, it *is* easier in many languages to learn many object words; on the other hand, children are also building up other categories of their language from the beginning. Developing categories

helps the child segment the other words they hear, and the relative growth of these categories depends on the characteristics of the language being learned and the way it is used in CDS.

Learning Grammar

Learning Morphology

Inflections change the grammatical meaning of words. Languages differ greatly in the amount and kind of inflectional morphology that they show. English has very little inflectional morphology (plural –*s* and past tense –*ed* are two examples). Free morphemes also mark grammatical meanings in languages. For instance, in English, *he* indicates the nominative masculine pronoun (subject argument), whereas *him* marks the accusative masculine pronoun (direct or indirect object argument). The English pronoun system shows irregularity: each pronoun in the English paradigm changes somewhat differently between nominative, accusative, and genitive (e.g., *he, him, his, we, us, our*). In addition, multiple meanings are encoded in one morpheme: number, case, and gender. Other languages encode meaning much more extensively and systematically with both inflectional and free morphemes. Inflections can mark either a local meaning (e.g., plural on a noun in English) or a relational meaning (e.g., plural on a verb to agree with the subject or object; a case marker on a noun as a function of its syntactic role). In either case, productive inflectional marking involves constituency: the ability to understand how plurality is marked on the whole noun phrase within the language being learned and how the noun phrase is coordinated with marking on the verb phrase. Thus, in some languages, determiners, adjectives, and nouns are all marked for plurality and/or case and/or gender, in others not; in some languages there is both subject and object marking on the verb and/or its auxiliaries and in others not.

Children may start by rote learning some frequent words together with their inflections while omitting nonsalient inflections (omission errors). Later, children may over-generalize and use incorrect forms (errors of commission). For instance, in English, children leave out tense markers (*It go there* for *It goes there*), over-regularize past tense marking (*goed* for *went*), and use incorrect forms (*I think her was crying for me*). In languages in which all words in a class carry an inflection, the child may initially use only one inflection and over-generalize it to all members of that class.

A central question is when children become productive with morphology. Researchers from different theoretical persuasions tend to interpret the data in predictably different ways. Those arguing for early, abstract knowledge point to the fact that children learning highly inflected languages produce words with inflectional marking from the beginning, usually accurately. Those who emphasize that the abstract representations underlying inflectional morphology are learned point to the potentially rote-learned nature of early inflectional marking and the limited range of items with which particular inflections are used.

There have been many different ways of defining productivity in naturalistic corpora: for instance, percentage provision in obligatory contexts, use of the same inflection on a number of members of the class, or use of different inflections on the same word. Inevitably, these measures yield different results, depending on how stringent the criteria are and how good the sampling is (Tomasello and Stahl, 2004). Methods are being developed that control for the range of forms in the child's lexicon and then compare the degree of productivity between the child's system at different developmental points and between the child's system and the adult's. For two children learning Spanish, a language with rich inflectional morphology on verbs, this method has shown that productivity develops even after the child can provide all the forms and that the children's range of inflectional provision on the same verbs as their parents, even when the range of inflections is controlled for, is still significantly different at age 2.6 (Aguado-Orea, 2004).

How quickly children learn the morphology of their language seems to depend on the following factors and the interaction among them: type frequency (how many different lexemes are inflected in the same way) and token frequency (the relative frequency of different surface forms) in the input; salience (can they be heard?); transparency of meaning (is the semantics accessible?); formal complexity (is one meaning or more encoded?); and the regularity and distributional consistency of the inflectional paradigm (see Slobin,1985; Peters, 1997 for summaries). In such languages as Turkish, in which suffixes are added in a consistent order and with consistent form to indicate local meaning (e.g., plural, location, possessive), children seem to be at least partially productive relatively early. Other languages present more difficulty – those in which the morphemes are: (1) portmanteau (more than one meaning is combined in one form; e.g., the nominative for 'water' in Russian is *vodá*; the dative plural form is *vodám*); (2) distributed (e.g., marked with a free form and an inflection, Serbo-Croatian locative markers); and/or (3) in complex and partial paradigms (e.g., Polish case endings).

One sure way to know that children have a productive system is to test for their ability to mark words that they have never heard before ('nonce' words). Otherwise, one can never know whether each word and its accompanying inflection have been learned from the input on a one-by-one basis. In Berko's classic 'wug test,' children are first introduced to a novel creature called *a wug*, and then another of these creatures is produced. The child is then asked, *Now we have two – ?* We know that children have productive plural morphology if they answer *wugs*.

This method can be extended to other word classes. For instance, experiments with novel verbs in English suggest that productive past tense morphology for verbs develops somewhat later than does plural morphology for nouns. Note that this is also a way of determining what type of category is represented in the child's system and its level of abstraction (for instance, action verbs only or all verbs or concrete nouns only or all nouns). Nonce experiments have not been conducted in many languages other than English, but their wider use would certainly help clarify some very complex issues in the learning of inflectional morphology. However, they are not easy to conduct with children younger than age 2, and children learning some languages may be partially productive before this age.

Because inflectional morphology can be extremely complex and children seem to demonstrate early sensitivity and productivity, it is often used as a testing ground for debates about whether children have very early abstract linguistic representations. An example is the extensive debate on past tense marking in English. Children start by producing a small number of forms (many of them irregular) and subsequently start to over-generalize the regular *–ed* marker to irregular verbs (e.g., '*runned*' rather than '*ran*'). Children concurrently use the correct form and may continue to over-regularize some forms for many years. Proponents of early abstract representations argue that, once children have learned the past tense rule, they seek to apply it to all new forms that they learn and stop doing so only when they learn the new irregular form that 'blocks' the application of the rule. This model is called the 'dual route model': one route for regulars and one for irregulars. The claim is that this learning occurs even in languages, unlike English, for which the rule is not the most frequent form (i.e., is a 'minority default' rule). Challenges to this position come from 'single-route models' that claim: (1) the developmental pattern of marking can be closely modeled in connectionist networks that have only one (not two) mechanisms (Plunkett and Juola, 1999); (2) these networks can also model the development of marking in languages in which the 'default' marking is not the most frequent; (3) detailed studies of children's marking in languages either with a 'default' (e.g., German plurals; Köpcke, 1998) or with no default (e.g., the Polish genitive; Dąbrowska, 2004) show a long process of development during which children over-generalize a variety of markers, depending on such factors as frequency and phonological similarity, before arriving at the adult system; and (4) the fact that over-generalizations continue for a long time after the child has learned the correct form does not suggest that there are two entirely separate processes involved (Maratsos, 2000).

As well as being a potential window on the nature of children's underlying syntactic abstractions, inflectional morphology can also provide a window on the development of children's underlying semantic representations. An example of this is the 'aspect before tense' debate. Here, the issue is whether children are initially sensitive only to the aspectual features of situations and their marking on verbs (e.g., punctual or durative events: *hit* vs. *singing)* or also to the temporal features (past vs. present events: *sit* vs. *sat*; Weist *et al.,* 2004). Because languages differ greatly in how these features are coded on verbs, cross-linguistic comparison of children's development has been essential in this debate. Methodological problems and the complexity and interrelatedness of the tense and aspect systems in different languages mean that we do not yet have a final answer to this issue.

Early Syntactic Development

Children's first utterances containing more than one word appear between 18 to 24 months of age. Almost all researchers agree that, in the early stages, some utterances are rote learned as a whole (e.g, *what's that?*), and others may be slot-and-frame patterns (e.g., *where's X gone?* or *more X*, where X denotes a range of referents). However, there is disagreement as to the extent of these low-scope patterns: what part they play, if any, in children's developing linguistic representations; and how utterances that are considered not to derive from them should be represented (Lieven *et al.*, 2003). For instance, when it is suggested that the child is using rules for the combination of underlying categories, the hypothesized categories can be very varied. One suggestion would be semantic categories; for instance 'agent' and 'action verb' to generate an utterance like *Lion swim*. However, a different researcher might analyze the same utterance in terms of syntactic relations, such as 'subject' and 'intransitive verb.' In the case of any particular utterance, these alternatives cannot be distinguished: A corpus of utterances (the larger, the more reliable) must be analyzed together with clear definitions of

productivity before it is possible to suggest whether the utterance is likely to have been generated productively and on the basis of what linguistic representations. As noted for morphology, naturalistic data never allow one to be completely certain that an utterance is productive.

Many of children's early utterances are missing features that are provided in adult speech: subjects are frequently omitted, as are function words (e.g., auxiliaries, complementizers, prepositions). Utterances often lack finiteness marking (e.g., no third person *–s* on main verbs in English; nonfinite for finite verbs in Dutch and German).

Children's utterances also reflect features of the input very closely. Thus, English-speaking children's early verbs are those that occur most frequently in the input, children mark verbs in accordance with the most frequent marking on those same verbs in the input, and they use verbs in the argument structures that their mothers use most frequently. In fact, whatever system one looks at, there are strong correlations with frequency in the input. This is not, however, the only factor affecting the child's use: as with phonology and semantics, complexity and salience also play a role, together with, most importantly, what the child wants to talk about.

None of these facts are in serious dispute in the literature; however, their interpretation certainly is. The major divide is between those who interpret them as indicating that the child has early abstract and, specifically, linguistic knowledge (for some researchers, innate) and those who argue that abstraction develops throughout the process of learning language. Clearly, even though children's utterances show a great deal of lexical specificity and partial marking, they could still reflect underlying abstract linguistic knowledge, limited by these four factors:

1. lexical learning of the forms of the language
2. learning the syntactic features particular to the language
3. processing constraints on production (e.g., a limit on the length of what the child can produce)
4. late biological maturation of some part of the system that underlies the abstraction.

Each of these factors is briefly considered below.

In addition, all theories need to account for the patterns of errors that children show. Those maintaining the existence of innate linguistic knowledge must answer this question: If children have innate grammar, why do they make errors that seem to reflect the lack of knowledge of fundamental linguistic systems, such as tense, agreement, and the provision of arguments? Those maintaining that abstractions are constructed must determine how the interaction between the child's current system and the input explains the pattern of errors.

In the case of (1) above, it is obvious that, whatever the underlying system, children have to learn the specific forms of their native language. The clear implication is that claims to underlying structure or the lack of it should not be made until it is established that the child is able to fully comprehend or produce the forms reflecting the hypothesized abstraction.

Within nativist linguistic approaches, one solution to the issue of how children use the hypothesized Universal Grammar (UG) to work out the syntactic particularities of their input language (2 above) is to suggest that they have to set a range of parameters as the result of hearing (a small number of) utterances. Within this literature, there are major debates on the number of parameters, whether or not they are initially set, and how many settings there can be (one, two, or more). An example is to hypothesize a 'head direction parameter' that determines how children identify the input language as being 'left' or 'right branching.' In principle, children then know how to order words across a range of phrasal structures (e.g., noun phrases, prepositional phrases, and verb phrases). Alternative constructivist accounts would suggest that children do this on the basis of the particular strings that they learn from the input, from which word-order patterns are subsequently abstracted in conjunction with the learning of word classes. This hypothesis would make it easier to account for the many instances of inconsistency of head direction in languages. The success of the Principles and Parameters approach in its own terms depends on the number of parameters being relatively small and on agreement among researchers as to how the specified settings can account for the range of relevant phenomena in language development (Fodor, 2001). At the present time, it is not clear that either of these conditions can be met.

In the case of processing constraints (3 above), there clearly are length constraints on what the child produces. Probably, the single most obvious aspect to the lay observer of children's developing language is that children's utterances get longer as they grow older. However, is this an output limitation imposed on a much fuller underlying abstract representation (Valian, 1991), or does it reflect what the child has actually learned up to this point? For instance, if children are processing the input by using already identified strings as the basis for building new strings, doing so would also yield increasingly long utterances over time (Elman *et al.*, 1997). In addition, the way in which patterns were added would mean that the distribution of marking would also change over

time. Research incorporating this type of learning mechanism has been successful at modeling children's patterns of finiteness marking not only in English but also in Dutch and Spanish (Gobet *et al.,* 2004).

An alternative linguistic nativist approach is to hypothesize that, although UG is innate, some parts of the system are on a later maturational timetable, which can account for some of the early error patterns shown by children (4 above). Two well-known accounts that use this approach are Borer and Wexler's attempt to explain the late use of full eventive passives by English-speaking children and Wexler's Agreement-Tense Omission Model (ATOM). In the first, the authors suggest that children are late with eventive passives because 'argument-linking chains' are a late biological maturation (Wexler, 2002). This theory cannot account for the fact that full passives emerge early in some languages (for instance in Sesotho; Demuth, 1992).

The ATOM model (of which the Optional Infinitive hypothesis for typically developing children and the Extended Optional Infinitive hypothesis for children with SLI are earlier versions; Wexler, 1998) predicts that, because pre-maturation children are subject to 'the unique checking constraint,' they will only be able to check either for tense or for agreement. This constraint is intended to account for patterns of incorrect finiteness marking found in young children's speech across a range of languages. One aspect of this model concerns agreement marking. Because, in the model, children can only check for either tense or agreement, a pattern of errors is predicted that, in English, sometimes results in non-nominative pronouns appearing in subject position (e.g., *me* for *I* errors or *her* for *she* errors). The only error predicted *not* to occur at levels greater than noise is the use of a non-nominative pronoun (e.g., *me, him, her, them*) with a finite verb that agrees with the subject (e.g., *Me am going, Her wants that*), because this would result from both tense and agreement having been checked. However, Pine *et al.* (2005) have shown that if agreement rates are broken down by the specific pronoun, errors with *her* in subject position and an agreeing verb occur too often to be disregarded as noise. One implication of this analysis is that it is important to analyze errors in terms of the particular form, rather than to sum across a category defined by the adult system, because the child's system may not be operating at this level of abstraction.

The other main theoretical approach to children's syntactic development argues that grammatical abstraction emerges from interaction between the input and innate learning mechanisms and that these are *not* restricted to the domain of syntax. Constructivist approaches maintain that children start by learning low-scope constructions based around specific words or morphemes. These constructions become more complex, and abstractions (for instance, of tense, agreement, subject, and transitive) build up. Initially, a child may have no understanding of the internal structure of a construction (e.g., *what's that?*), but uses it as a whole with a specific meaning. As development proceeds, functional distributional analysis based on the relation between a form and (child-identified) functions leads to representations of constructions developing internal structure. Patterns of relationships build up between constructions and their parts, in a process of increasing complexity and schematization (Table 1).

One critical aspect of this theory is that, once a slot develops in a construction, it will be paired with a meaning: thus, a 'noun phrase' slot always denotes a referent. In time, the child will learn to refer in increasingly complex ways; for instance, in English with a range of determiners and adjectives. Constituency thus develops hand in hand with the function of the constituent in the construction and becomes increasingly schematic (Dąbrowska, 2004).

Tomasello (2003) provided the most comprehensive statement of this position to date. He suggested that, rather than having to postulate innate syntactic categories or linking rules, it is possible to account for the development of abstract categories and constructions on the basis of general communicative and cognitive processes, such as intention reading, analogy making, and distributional analysis. The critical aspect of this approach is that utterances have meanings as a whole and that, for children to be able to work out

Table 1 Examples of different types of constructions and their meanings

Form	*Meaning*	*Level of abstraction*
What's that?	Requesting a name	Rote-learned
It's a <Noun>[1]	Naming <referent>	Slot-and-frame
Double object construction: <Noun> <Verb> <Noun> <Noun> *Mary gave John the ball*	Agent-Verb of transfer-Recipient-Transferred Object	Schematic

[1]<> denotes an open slot.

what role a word or inflection is playing in an utterance, they have to know, first, what the whole utterance means and then, in a process that Tomasello calls 'blame assignment,' identify the part of the meaning for which the particular form is responsible. In this approach, abstract constructions develop through analogies based on semantic similarity between the meanings of different item-based constructions; for instance, *A hits B*, *X loves Y*, *C pushes D* may give rise to the transitive construction and its syntactic roles.

Tomasello's (1992) 'verb island hypothesis' was an early example of such an approach. He claimed that children initially build up constructions around individual verbs, rather than having more general and abstract categories, such as subject, direct object, and transitive verb, from the beginning. Thus, a child who knows the verb *hit* may have a slot before the verb for the 'hitter' and one after for the 'hittee' without having the abstract representation of a transitive construction. A wide range of production experiments that test English-speaking children's ability to use novel verbs in transitive constructions supports this claim (see Tomasello, 2003). Although most children aged just over 2 years old are already using verbs they know in some two-argument constructions, the ability to demonstrate this knowledge with a verb they have never heard before develops between the ages of 2 and 3 and can be affected strongly by the experimental method employed.

The verb island hypothesis has been challenged by studies using the preferential looking paradigm. They suggest that children can discriminate between a transitive utterance that matches a scene they are looking at and one in which the argument roles are reversed at younger ages than those shown for success in production experiments. Some researchers have suggested that this indicates that children are innately equipped either with a notion of subject or with rules that link the notion of agent to subject. However, it seems difficult to argue for an innate notion of subject, because subject is not a universally identical category across languages (Croft, 2001).

Other researchers, less contentiously, suggest that children are already sensitive to the number of arguments in an utterance and their relationship to a causative scene with two actors and that this allows them to make the preferential-looking discrimination (see the debate involving Fisher and Abbot-Smith and Tomasello in *Cognition*, following Tomasello's (2000) article). Because children are sensitive pre-verbally to causation and many already know how to name people and objects by the time they take part in preferential-looking studies, this finding seems plausible – particularly because there seems to be a correlation between the child's vocabulary level and success in these tasks.

Children's avoidance of using ungrammatical transitives and intransitives in production experiments also suggests earlier sensitivity to some aspects of verb argument structure. In addition, experiments show that English-speaking children manage to use the transitive construction with a novel verb somewhat earlier when case-marked pronouns, rather than nouns, are used as arguments. It may be that languages, such as German, with relatively clear case-marking of argument roles, also assist children in schematizing these roles somewhat earlier. Pine *et al.* (1998) have argued for patterns building up not only around verbs but also around other high-frequency markers, such as *I* + *Verb* or *Verb* + *it*. Thus, the child, as well as building up a 'transitive verb' category, would, at the same time, be learning that *I* and other frequent referents in first position can be in a range of semantic roles in relation to the verb. Childers and Tomasello found supporting evidence for this suggestion (see Tomasello, 2003). Thus, more recent constructivist accounts suggest that the abstract transitive construction builds up from a number of different sources: pre-verbal knowledge of causality, the ability to match arguments in an utterance to referents in the environment, the development of other lexically specific constructions around pronouns, and high-frequency items. The combination of these factors may result in the child being able to perform correctly in some tasks using less fully specified representations than would be required for others (e.g., using a novel verb that has never been heard before in a transitive construction).

From a constructivist perspective, there is initially no necessary relation in the child's system between one construction and another; for instance, between a declarative and a *wh*-question that would be related in the adult system as would be maintained in any analysis based on one or other generativist theory of island constraints (e.g., movement, gaps or checking). However, such relationships do develop. This can be seen from children's increasing conversational flexibility as they question previously mentioned constituents, manipulate topic and focus, and, in English, expand contracted auxiliaries, insert DO-support when rephrasing a previous utterance, and so on. Many experimental studies also show that children become increasingly able to hear a nonce form in one construction (for instance, an active) and to transform it into another (a passive). For linguistic nativists, this should be a relatively automatic outcome of children's abstract linguistic representations.

However, within constructivist approaches, these relations are seen as being learned. This learning

occurs in part through the identification of the same meaning-based categories in different constructions and in part through relating the overall meanings of constructions to each other. For instance, the simple active statement, *I want a banana,* communicates something different from the cleft: *It's a banana I want*. Children's over-generalizations of constructions are explained relatively easily within this framework: For instance, when a child says *Don't giggle me* (meaning *Don't make me giggle*), she has analogized *giggle* as having causative meaning and therefore placed it into a causative transitive construction. There are, however, considerable problems both for linguistic nativists and construction-based accounts in explaining how children cut back from such over-generalizations (Bowerman, 1996; Tomasello, 2003).

Later Syntactic Development

As children's language develops, major changes occur at both clausal and sentential levels. Noun phrases and verb phrases increase in complexity: for instance, although English-speaking children use *a* and *the* early on, it may take many years before they can operate with more complex determiners and their scope (*some*, *every,* and *each*). Likewise, children may be able to use some auxiliaries (e.g., *I'm X-ing*, *Can I Y?*) and verb complement structures (e.g., *I want to play out*) from early on, but the full development of complex verb phrases takes considerable time. At the sentential level, although English-speaking children produce many questions between the ages of 2 and 3, initially most of these are with contracted copulas (*What's that?*, *Where's X?*) or contracted auxiliaries (*What's Mummy X-ing?*, *Where's Y gone?*). Some questions with subject-auxiliary inversion (*Can I X?*, *Are you Y-ing?*) also start appearing during this period, but it is not until approaching age 3 and older that children start to show considerable flexibility in the range of subjects, auxiliaries, and *wh*-words that they can use in inverted *yes/no*- and *wh*- questions. Clause coordination and subordination (e.g., complement structures, causatives, relative clauses) also develop over a considerable time period – both when these structures start to appear and when they increase in internal complexity.

In principle, it is rather difficult for linguistic nativist theories to explain this slow and somewhat patchy development of complex syntax and the accompanying errors that children make. Once the lexical forms are learned, why should children make uninversion errors (*Where she is going?*) or find some embedded relatives easier than others? Linguistic nativist explanations of children's later syntactic development all start from the claim that children could not learn the abstract basis of these constructions because the constituents involved could not be abstracted from a surface analysis of the input. Errors are accounted for in two ways: (1) they arise from the particularities of the language being learned and not from the absence of the abstract representations of Universal Grammar; or (2) children have full competence and errors arise from faulty methodologies (see O'Grady, 1997, for a detailed coverage of explanations from within a broadly UG framework). In constructivist accounts, complex constructions, such as *wh*-questions or relative clauses, begin as item-based or with simpler structures that then build toward greater complexity and schematicity (Tomasello, 2003).

Uninversion errors in English-speaking children's *wh*-questions have given rise to several different linguistic nativist accounts. For example, de Villiers has proposed that questioning arguments is easier than questioning adjuncts. An alternative account, related to the particularities of English (Santelmann *et al.,* 2002), suggested that the errors in children's inversion are explained by the fact that only main verb BE can invert (unlike in German where all main verbs can invert in questions) and DO-support is also unique. Recent constructivist accounts have explained inversion errors in terms of the relative frequency of *wh* + auxiliary combinations in the input. Thus, Rowland and Pine (2000) suggested that children are learning correct inversion from high-frequency combinations of a *wh*-word plus a specific form of the auxiliary and that uninversion errors occur when they have not yet learned the correct combination and produce a 'groping pattern' by putting together a known *wh*-word and a known declarative.

However, the overriding argument, from the linguistic nativist position, against the possibility of learning syntax is that children can deal with complex constructions that could not have been learned from the input. An example is extraction from embedded clauses in *wh*-questions: *Who do you think John likes?* In UG accounts (see, for instance, Crain and Thornton, 1998; Drozd, 2004), children operate with abstract formalisms that relate the object of the embedded clause first to the object of *you think* and from there to the position at the front of the sentence. These formalisms could not, it is argued, be learned from the sequential probabilities of strings on the surface of the input; in addition, the structures are rare to nonexistent in the input and therefore could not be learned as a whole. Whether this latter point is the case awaits much denser sampling of the input, but it should be noted that this type of extraction is not possible in German and that a much wider range of extraction occurs in Italian. This considerable cross-linguistic variation therefore requires, in UG

terms, a complex parametric account and, from the point of view of the language learner, close attention to the particularities of the language.

From a constructivist point of view, children solve these tasks by building up their knowledge of constituency and using already established form-meaning mappings to identify potential constituents in novel constructions. Thus, although children may not have had previous experience with the type of sentence in the Crain and Thornton experiments, they will already know constructions that identify the matrix *wh*-clause as a constituent asking for an act of reference, thus allowing them to attempt to identify the referent of the embedded relative. For instance, English-speaking children's early relative clauses tend not to be restricted (i.e., adding more information about a referent: *The cow that the dog bit ran away*), but presentational (i.e., introducing a topic: *That's the cow that goes there*). Presentational relatives do not involve embedding one clause within another; the main clause is usually a copula (one of the earliest constructions that English-speaking children learn), and the relative clause is not presupposed information but contains the new information to be conveyed. Diessel and Tomasello suggested that all these factors make these constructions easier to produce and also make them the basis for the development of more complex relative clause structures. Diessel (2003) has similar accounts for the development of a number of other constructions in English-speaking children's development. The idea that already existing constituents can be used to work out the meaning of a novel construction is supported by a modeling study (Morris *et al.,* 2000) in which a connectionist network was shown to generalize correctly to previously unlearned constructions on the basis of having learned a variety of simpler constructions.

Developing Pragmatic and Metalinguistic Skills

Children are immersed in communicative interaction with others from infancy, but there are many aspects that take years to develop. Coordinating reference across speaker and listener roles involves manipulating and understanding given and new information, deictic and anaphoric pronouns, temporal information, and so on. In most of the cultures studied, these skills develop initially in conversations with others, and children's attempts to narrate sequences of events start to develop slightly later. Both conversation and narrative require the ability to ground utterances and to refer to referents in ways that require complex cognitive, communicative, and linguistic skills. Children also have to learn to adjust their speech to the genre of the task (conversation, personal narrative, reporting, arguing) and to fit the social context in which they are communicating (for instance, levels of politeness and formality; Slobin *et al.,* 1996).

Conversation

Early conversations with young children typically show a great deal of scaffolding. Thus, background and setting and the identification of referents are provided by the child's interlocutor. Around the age of 2, children are more fluent participants when the conversation is highly 'scripted' around routine activities and when they can join in a multiparty conversation at moments that they choose. Between the ages of 2.6 to 3.0, the first, more subtle, discourse particles start to be used (for instance, in English, *now* and *just*), but their range and presence continue developing up to and beyond the school years. Topic continuity in terms of responding to questions and maintaining the topic over a number of turns also improves radically between the ages of 2 and 4. Clarification requests by children and children's responses to them show clear development as well: Children respond to clarification requests with more nuanced answers as they get older, and they also make more of them, though even at the age of 7, children in a referential communication task do not make clarification requests either as much or in as focused a way as is required by the task (Lloyd, 1991, cited in Tomasello, 2003). Children's rhetorical skills develop throughout the school years, with the ability to bring arguments to bear, to return to a conversational topic, and to introduce new topics in an appropriate way all taking many years to perfect.

Narrative

Children's early narratives tend to be a set of short utterances with no attempt at cohesion among them. For a child to be able to narrate an event to another who was not present requires taking into account the hearer's lack of knowledge of the timing of the event, its sequence, and the participants. This is a complex task in itself, and children also have to learn to coordinate sequences of utterances to keep track of referents and events such that, for instance, already referred-to participants are referred to in one way, and novel ones in another. Two major summaries of studies that have looked at the development of these skills within and across cultures and languages are Berman and Slobin (1994) and Hickmann (2003).

Once children start being able to produce some coordination between utterances in a narrative, they often use very repetitive devices: for instance, maintaining reference to the same referent as thematic

subject throughout the story (e.g., using *He* at the beginning of every utterance) or using *and then* at the beginning of every utterance to coordinate the action. Clearly, the development of narrative abilities is based on the development of a range of skills: Children must be able to take into account the perspective of the listener in terms of what the listener does and does not know. They must also know the linguistic devices of the language that coordinate clauses within utterances (e.g., temporal sequencing, subordination) and across utterances (e.g., the use of pronouns for anaphoric reference), and they must know how their language encodes perspectives; for instance, in English, manner of motion tends to be encoded in the verb, whereas in Spanish it tends to be encoded in adverbials.

Metalinguistic Skills

Children gradually become able to consciously reflect on language, to correct their own and others' utterances, and to notice features of language at all levels: Rhyming relationships, the explanation of metaphors and idioms, and grammaticality judgments. Karmiloff-Smith (1994) has incorporated language development and metalinguistic awareness into her general theory of the development of cognitive skills. In this approach, the early period of language learning is concerned with skill automatization: children have to learn the forms of the language and to use them relatively effortlessly. Once language becomes relatively fluent and automatic, it can become an object of reflection.

Karmiloff-Smith also discussed children's increasing awareness of the plurifunctionality of forms. She sees metacognition as going through several stages in which initially awareness is implicit and only later, as a result of 'representational redescription,' does it become available to conscious reflection.

For many children, the process of moving to conscious awareness of the forms of language and how they fit together is almost certainly aided by learning to read and write (Ravid and Tolchinsky, 2002). Miller and Weinert (1998) give a detailed cross-linguistic account of the ways in which oral and written language vary and the implications of this variation for theories of linguistic competence. For instance, they argue that many of the structures that linguists are concerned with can only be found in written texts and thus that children only learn them during the school years. Competence is, therefore, still developing, rather than being either innate or language development being 'all over by 3.0.' How rapidly these skills develop and how far children get with them are related to the complexity of the language that they hear at home and at school and to levels of literacy (Hart & Risley, 1995; Huttenlocher *et al.*, 2002).

Clearly, literacy cannot be the only way in which speakers become aware of language as an object of reflection. Pre-literate children already show some metacognitive awareness of aspects of language as indicated by Karmiloff-Smith; for instance, they can sometimes correct ungrammatical utterances in naturalistic and experimental situations. There must be many aspects of oral interactions that also afford this ability; for instance, argumentation, speech making, traditions of oral narrative, and the learning of complex ritual language are all likely to be media through which speakers and listeners come to be able to reflect on linguistic structure and its meaning. The important point is that these genres may contain much more complex language than children are likely to learn early on. Thus, language competence, often in the structures in which linguists are most interested, continues to develop over many years.

Atypical Development

Some children grow up in situations that are radically different from those in which typically developing children learn an oral language. The study of these children's atypical development often throws light on the processes involved in language learning and their inter-relationships, though caution should be exercised because one has to see the child and his or her situation as a whole. It is therefore impossible to treat these cases as controlled experiments in which only one variable is changed.

First, from the study of children who grow up from birth in a home where the language used is one of the sign languages of the world, it is clear that these children go through the normal stages of language development on the same timetable as children learning oral languages. Sign languages differ from oral languages not only in the obvious dimension of the medium but also, as a result, in the degree of simultaneity of signs and in the ways in which the face and hands are used grammatically and suprasegmentally. Thus, comparisons between the ways in which children learn the morphology, syntax, and pragmatics of oral and sign language can be extremely illuminating.

There are two groups of children whose nontypical development has been of particular interest to the study of language development: children whose language seems relatively poor by comparison with their nonverbal cognitive levels (children with SLI) and children with Williams's syndrome, whose language seems relatively advanced compared to their very low level of nonverbal cognitive development. Researchers coming from a more linguistic nativist background have claimed that these two syndromes

are evidence for an innate basis to grammatical development (compromised in SLI, preserved in Williams's syndrome) that is separate from more general cognitive development. In the case of SLI, the issue is the nature of the biological compromise. Most children with SLI have a range of problems with language, some of which may be more typical earlier and some later. Articulatory and phonological problems, together with delayed inflectional marking, are typical early problems that often resolve, leaving apparent grammar-only problems later on. There is certainly no agreement in the field that the pattern of impairment in SLI children indicates the presence of an innate module for syntax (Leonard, 1998), though some have made claims to this effect (Rice and Wexler, 1996).

In the case of Williams's syndrome, it is clear that these children also suffer from a wide range of problems with language and that none of them ever reach the language levels of age-matched, typically developing controls. Karmiloff-Smith *et al.* (2003) emphasized a developmental model for these children that suggests that, although language development is atypical from the outset, the problems that children demonstrate change with development and depend on the strategies for producing language that they have developed at earlier stages in the absence of the processes available to typically developing children. Here, too, then, there seems no clear evidence for a specific innate module for syntax.

Learnability and Constituency

The field of language acquisition has been dominated until recently by the so-called learnability issue; namely, that children cannot learn correct syntax from only positive evidence (the so-called poverty of the stimulus argument). Chomsky's various versions of a postulated innate Universal Grammar were attempts to provide these constraints, and the various nativist proposals outlined above aimed to show how these constraints explain different aspects of language development.

Several solutions have been proposed, all of which involve changing the premises of the learnability problem in one way or the other, usually to make the generation of sentences in the language probabilistic, rather than all or nothing (see MacWhinney, 2004 and commentaries). Although the leap from the higher reaches of statistical theory to the details of children's language development is rather large, this solution actually accords quite well both with constructivist approaches to language learning that emphasize the learning of the distributional probabilities in the input and to various attempts to model computationally how distributional probabilities might lead to phrasal learning (Elman *et al.*, 1997). Despite not incorporating any model of semantics, these attempts have been quite successful at detecting distributional regularities in the target language that could form one basis for the establishment of categories and of syntagmatic learning (however, see Chang, 2002).

However, language learning must go beyond the mere learning of immediately sequential probabilities between words. Simply calculating the transitional probability between one word and the next will ignore constituency and produce ungrammatical sentences. This was the basis for Chomsky's 1957 attack on the behaviorist theory of language development. Sentences involve constituents; the subject noun phrase, however long it is, agrees with the main verb, regardless of how many subordinate clauses occur in between; the relationship between an inverted question and its declarative version involves the relationship between the *wh*-word and the whole questioned constituent, independent of the length of the constituent or what occurs between it and the main verb. Thus in *The boy wearing a green hat who I saw yesterday in the park is very nice*, the subject NP (*The boy*) must agree with the main verb (*is*). The relationship between *Is the boy who is wearing a green hat nice?* and *The boy who is wearing a green hat is nice* is that the first *Is* in the question parallels the second *is* in the declarative; otherwise, we would have *Is the boy who wearing a green hat is nice?*

From a constructivist point of view, the crucial extra ingredient is that children are intention readers and meaning makers and that they learn about constituency by building up increasingly complex paradigmatic slots that connect form to meaning. Thus, a child who has heard and used literally hundreds of thousands of copula sentences of the form – *The boy's nice*; *The bus is red; They're the ones I want; They are really really good*; *The ones over there are the ones I want*; *The boy wearing a green hat is nice* – will know that *the boy wearing a green hat* refers to a particular object or person. This learned structural knowledge of constituency is then used to parse and produce the more complex embedded structures (see Lewis and Elman, 2001, for an attempt to model the learning of distance dependencies and constituency).

Conclusion

Humans have developed language and it has a biological basis. Yet, precisely what this basis is (or, more likely, bases are) still needs a great deal of precise specification. For instance, localization of language in the brains of adults is much more determined

than it is in children: Children who have early damage to the areas of the brain that subsume language in adults can learn language to normal levels and will have language represented in other parts of the brain as adults. Thus, it seems that language modularity, and the relative independence of syntax and semantics that characterizes the results of some experiments with adults, may be an outcome of development, rather than being pre-given. Chomsky's insistence on a separation between syntax and meaning, as well as the claim that syntax is essentially unlearnable and therefore that there must be an innate module for grammar, has ultimately had a negative impact on the scientific study of how children learn to talk. However, many of the issues raised by researchers working within this tradition, especially with regard to the development of more complex syntax, present challenges that have not yet been met fully by those who suggest that, whatever else is innate, syntax is learned through the application of general cognitive skills to the communicative tasks set for children by those around them.

Both positions face a number of challenges. In both, there is a need for greater specification of the processes involved so that proposals can be tested empirically. The challenge to linguistic nativists is to specify in advance the postulated performance limitations and late maturing processes precisely enough so that, rather than invoking them on an *ad hoc* basis to deal with counter-evidence, they can be used to make falsifiable predictions. In addition, these accounts often lack any notion of development: how does what a child is doing at one stage affect what happens next? Finally, because there are input effects at all levels of language acquisition, such theories tend to under-theorize the role of the input, not just in the quantity required by the child but also in the ways in which its precise characteristics are reflected in the child's system and how this interacts with any hypothesized innate knowledge.

The challenge to constructivists is to specify more precisely the ways in which the cognitive and communicative development of the child and the distributional information in the input interact to generate the patterns of learning and of errors and to test this with a wider range of languages and a wider range of input densities and conditions. There is a long way to go, even in the study of English language learning, to say nothing of other languages, in specifying the precise ways in which children expand their inventory of constructions and how they become increasingly schematic.

Children are creative communicators from the start. The question is the basis of this creativity, its scope, and how it develops.

See also: Bilingual Language Development: Early Years; Brown, Roger William (b. 1925); Bruner, Jerome (b. 1916); CHILDES Database; Chomsky, Noam (b. 1928); Cognitive Linguistics; Constituent Structure; Corpus Linguistics; Developmental Relationship between Language and Cognition; Discourse, Narrative and Pragmatic Development; Formal Models and Language Acquisition; Human Language Processing: Connectionist Models; Infancy: Sensitivity to Linguistic Form; Innate Knowledge; Langacker, Ronald (b. 1942); Language Acquisition Research Methods; Language Development in Blind Children; Language Development in Deaf Children with Hearing Parents; Language Development in School-Age Children, Adolescents, and Adults; Language Development: Morphology; Long-Distance Dependencies; Meaning: Development; Modularity of Mind and Language; Oral Traditions and Spoken Discourse; Phonological Awareness and Literacy; Pragmatics: Optimality Theory; Principles and Parameters Framework of Generative Grammar; Productivity; Quine, Willard van Orman (1908–2000); Relative Clauses; Sign Language: Acquisition; Slobin, Dan Isaac (b. 1939); Social-Cognitive Basis of Language Development; Specific Language Impairment; Syntactic Development; Syntax-Pragmatics Interface: Overview; Variation in First Language Acquisition; Williams Syndrome.

Bibliography

Aguado-Orea J (2004). *The acquisition of morpho-syntax in Spanish: implications for current theories of development*, Ph.D diss., University of Nottingham.

Bates E (1976). *Language and context: the acquisition of pragmatics.* New York: Academic Press.

Bates E, Bretherton I & Snyder L (1988). *From first words to grammar: individual differences and dissociable mechanisms.* Cambridge: Cambridge University Press.

Berman R & Slobin D (1994). *Relating events in narrative: a crosslinguistic developmental study.* Hillsdale, NJ: Erlbaum.

Bernhardt B & Stemberger J S (1998). *Handbook of phonological development.* San Diego: Academic Press.

Bowerman M (1988). 'The "no negative evidence" problem: how do children avoid constructing an over-general grammar?' In Hawkins J A (ed.) *Explaining language universals.* Oxford: Blackwell.

Bowerman M & Choi S (2001). 'Shaping meanings for language: universal and language-specific in the acquisition of spatial semantic categories.' In Bowerman M & Levinson S (eds.) *Language acquisition and conceptual development.* New York: Cambridge University Press. 475–511.

Carey S & Bartlett E (1978). 'Acquiring a single new word.' *Paper and Reports on Child Language Development 15,* 17–29.

Carpenter M, Nagell K & Tomasello M (1998). 'Social cognition, joint attention and communicative competence from 9–15 months of age.' *Monographs of the Society for Research in Child Development 255.*

Chang F (2002). 'Symbolically speaking: a connectionist model of sentence production.' *Cognitive Science 26(5),* 609–651.
Clark E (2003). *First language acquisition.* Cambridge: Cambridge University Press.
Crain S & Thornton R (1998). *Investigations in Universal Grammar: a guide to experiments on the acquisition of syntax.* Cambridge, MA: MIT Press.
Croft W (2001). *Radical construction grammar: syntactic theory in typological perspective.* Oxford: Oxford University Press.
Dąbrowska E (2004). *Language, mind and brain.* Washington, DC: Georgetown University Press.
Demuth K (1992). 'The acquisition of Sesotho.' In Slobin D I (ed.) *The crosslinguistic study of language acquisition,* vol. 3. Hillsdale, NJ: Lawrence Erlbaum. 557–638.
Diessel H (2003). *The acquisition of complex sentences in English.* Cambridge: Cambridge University Press.
Drozd K (2004). 'Learnability and linguistic performance (plus commentaries).' *Journal of Child Language 31,* 431–457.
Elman J, Bates E, Johnson M, Karmiloff-Smith A, Parisi D & Plunkett K (1997). *Rethinking innateness: a connectionist perspective on development.* Cambridge, MA: MIT Press.
Fenson L, Dale P, Reznick J S, Bates E, Thal D & Pethick S (1994). 'Variability in early communicative development.' *Monographs of the Society for Research in Child Development 242(59), 5.*
Fernald A, Pinto J, Swingley D, Weinberg A & McRoberts G (1998). 'Rapid gains in speed of verbal processing by infants in the second year.' *Psychological Science 9,* 28–31.
Fodor J (2001). 'Setting syntactic parameters.' In Baltin M & Collins C (eds.) *The handbook of contemporary syntactic theory.* Oxford: Blackwell. 730–767.
Gentner D (1982). 'Why nouns are learned before verbs: Linguistic relativity versus natural partitioning.' In Kuczaj S (ed.) *Language Development,* vol. 2. Hillsdale, NJ: Erlbaum. 301–333.
Gobet F, Freudenthal D & Pine J M (2004). 'Modelling syntactic development in a cross-linguistic context.' In Sakas W G (ed.) *Proceedings of the First COLING Workshop on Psycho-computational Models of Human Language Acquisition.* 53–60.
Gomez R L, Gerken L & Schvaneveldt R W (2000). 'The basis of transfer in artificial grammar learning.' *Memory & Cognition 28(2),* 253–263.
Hart B & Risley T (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: H. Paul Brookes.
Hickmann M (2003). *Children's discourse: person, space and time across languages.* Cambridge: Cambridge University Press.
Huttenlocher J, Vasilyeva M, Cymerman E & Levine S (2002). 'Language input and child syntax.' *Cognitive Psychology 45(3),* 337–374.
Jusczyk P (1997). *The discovery of spoken language.* Cambridge, MA: MIT Press.
Karmiloff-Smith A (1994). 'Precis of beyond modularity: a developmental perspective on cognitive science (with peer commentary).' *Behavioral and Brain Sciences 17(4),* 693–706.
Karmiloff-Smith A, Brown J H, Grice S & Paterson S (2003). 'Dethroning the myth: cognitive dissociations and innate modularity in Williams's syndrome.' *Developmental Neuropsychology 23(1 & 2),* 229–244.
Köpcke K (1998). 'The acquisition of plural marking in English and German revisited: schemata versus rules.' *Journal of Child Language 25,* 293–319.
Leonard L (1998). *Children with specific language impairment.* Cambridge, MA: MIT Press.
Lewis J D & Elman J (2001). 'A connectionist investigation of linguistic arguments from poverty of the stimulus: learning the unlearnable.' In Moore J D & Stenning K (eds.) *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Erlbaum.
Lieven E (1997). 'Variation in a crosslinguistic context.' In Slobin D I (ed.) *The crosslinguistic study of language acquisition,* vol. 5. Hillsdale, NJ: Lawrence Erlbaum. 199–263.
Lieven E, Behrens H, Speares J & Tomasello M (2003). 'Early syntactic creativity: a usage-based approach.' *Journal of Child Language 30,* 333–370.
MacWhinney B (2004). 'A multiple process solution to the logical problem of language acquisition (plus commentaries).' *Journal of Child Language 31,* 883–914.
Mandler J (2000). 'Perceptual and conceptual processes in infancy.' *Journal of Cognition and Development 1,* 3–36.
Maratsos M (2000). 'More overgeneralisations after all.' *Journal of Child Language 28,* 35–54.
Markman E, Wasow J & Hansen M (2003). 'Use of the mutual exclusivity assumption by young word learners.' *Cognitive Psychology 47(2),* 241–275.
Miller J & Weinert R (1998). *Spontaneous spoken language: syntax and discourse.* Oxford: Clarendon.
Morris W, Cottrell G & Elman J (2000). 'A connectionist simulation of the empirical acquisition of grammatical relations.' In Wermter S & Sun R (eds.) *Hybrid neural symbolic integration.* Berlin: Springer Verlag.
O'Grady W (1997). *Syntactic development.* Chicago: University of Chicago Press.
Peters A (1997). 'Language typology, prosody and the acquisition of grammatical morphemes.' In Slobin D I (ed.) *The crosslinguistic study of language acquisition,* vol. 5. Hillsdale, NJ: Lawrence Erlbaum. 135–197.
Pine J, Lieven E & Rowland C (1998). 'Comparing different models of the development of the verb category.' *Linguistics 36,* 4–40.
Pine J, Rowland C, Lieven E & Theakston A (2005). 'Testing the Agreement/Tense Omission Model: why the data on children's use of non-nominative 3psg subjects count against the ATOM.' *Journal of Child Language 32,* 2.
Plunkett K & Juola P (1999). 'A connectionist model of English past tense and plural morphology.' *Cognitive Science 23(4),* 463–490.

Ravid D & Tolchinsky L (2002). 'Developing linguistic literacy: a comprehensive model (plus commentaries).' *Journal of Child Language 29,* 417–447.
Rice M & Wexler K (1996). 'Towards tense as a clinical marker of specific language impairment in English-speaking children.' *Journal of Speech and Hearing Research 41,* 1412–1431.
Rowland C & Pine J (2000). 'Subject-auxiliary inversion errors and *wh*-question acquisition.' *Journal of Child Language 27(1),* 157–181.
Santelmann L, Berk S, Austin J, Somashekar S & Lust B (2002). 'Continuity and development in the acquisition of yes/no question: dissociating movement and inflection.' *Journal of Child Language 29,* 813–842.
Slobin D I (1985). 'Crosslinguistic evidence for the language-making capacity.' In Slobin D I (ed.) *The crosslinguistic study of language acquisition,* vol. 2. Hillsdale, NJ: Lawrence Erlbaum. 1157–1256.
Slobin D I, Gerhardt J, Kyrattzis A & Guo J (1996). *Social interaction, social context and language: essays in honor of Susan Ervin-Tripp*. Mahwah, NJ: LEA.
Tomasello M (1992). *First verbs: a case study of early grammatical development*. New York: Cambridge University Press.
Tomasello M (2000). 'Do young children have adult syntactic competence?' *Cognition 74,* 209–253.
Tomasello M (2003). *Constructing a language*. Cambridge, MA: Harvard University Press.
Tomasello M & Stahl D (2004). 'Sampling children's spontaneous speech: How much is enough?' *Journal of Child Language 31,* 101–121.
Valian V (1991). 'Syntactic subjects in the early speech of American and Italian children.' *Cognition 40,* 21–81.
Vihman M (1996). *Phonological development*. Oxford: Blackwell.
Weist R, Pawlak A & Carapella J (2004). 'Syntactic-semantic interface in the acquisition of verb morphology.' *Journal of Child Language 31(1),* 31–60.
Wexler K (1998). 'Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage.' *Lingua 106,* 23–79.
Wexler K (2002). 'Lenneberg's dream: learning, normal language development and specific language impairment.' In Schaffer J & Levy Y (eds.) *Language competence across populations: towards a definition of specific language impairment*. Mahwah, NJ: Lawrence Erlbaum.

Language Development: Pre-scientific Studies

R N Campbell, University of Stirling, Stirling, UK

The scientific study of language development may be characterized as involving collection of dated observations linked to general development; the systematic study of such collections by means of type and token counts, tabulation, and so forth; and the development of explanatory accounts grounded in such collections. By the beginning of the 20th century, this sort of apparatus was well established, and works in book form such as Major (1906) and Stern and Stern (1907) presented what are undeniably scientific treatments. However, the foundation of the journal *Pedagogical Seminary* by G. Stanley Hall at Clark University in 1891 provides a more accurate boundary mark: from its foundation onward, the journal featured systematic accounts of aspects of language development.

Early Studies

Reported empirical investigation of language development can probably be dated from the Mughal Akbar the Great's isolation experiment (c. 1578) in which a large number of children were isolated at birth in his palace at Fatehpur Sikri for a period of four years and cared for by unspeaking nurses. According to the three contemporary reports of this experiment, the children did not speak, but may have vocalized. The most explicit account is given in the *Akbarnama*, a life of Akbar prepared by his minister, Sheikh Abu'l Fazl, in the form of an exquisitely illuminated manuscript. Three such manuscripts were produced in the royal workshops. One is located in the Victoria & Albert Museum in London; a second is split between the Chester Beatty Library, Dublin and the British Library; the third is dispersed, with many folios held privately. For printed versions, see Blochmann (1877–86) and Beveridge (1902–1939), who reports the outcome as: "No cry came from that house of silence, nor was any speech heard there. In spite of their four years they had no part of the talisman of speech, and nothing came out except the noise of the dumb." Similar experiments are said to have been carried out at earlier times by the Egyptian Pharaoh Psamtik I in the 7th century B.C., by Emperor Frederick II of Sicily in the 13th century, and by King James IV of Scotland in the late 15th century. However, there are good reasons to doubt the authenticity of all three. For reviews of these early experiments, see Tylor (1878: 79–81), Campbell and Grieve (1982), and Bonvillian *et al.* (1997). It is perhaps remarkable that the first attempt to study language

Ravid D & Tolchinsky L (2002). 'Developing linguistic literacy: a comprehensive model (plus commentaries).' *Journal of Child Language 29,* 417–447.
Rice M & Wexler K (1996). 'Towards tense as a clinical marker of specific language impairment in English-speaking children.' *Journal of Speech and Hearing Research 41,* 1412–1431.
Rowland C & Pine J (2000). 'Subject-auxiliary inversion errors and *wh*-question acquisition.' *Journal of Child Language 27(1),* 157–181.
Santelmann L, Berk S, Austin J, Somashekar S & Lust B (2002). 'Continuity and development in the acquisition of yes/no question: dissociating movement and inflection.' *Journal of Child Language 29,* 813–842.
Slobin D I (1985). 'Crosslinguistic evidence for the language-making capacity.' In Slobin D I (ed.) *The crosslinguistic study of language acquisition,* vol. 2. Hillsdale, NJ: Lawrence Erlbaum. 1157–1256.
Slobin D I, Gerhardt J, Kyrattzis A & Guo J (1996). *Social interaction, social context and language: essays in honor of Susan Ervin-Tripp*. Mahwah, NJ: LEA.
Tomasello M (1992). *First verbs: a case study of early grammatical development*. New York: Cambridge University Press.
Tomasello M (2000). 'Do young children have adult syntactic competence?' *Cognition 74,* 209–253.
Tomasello M (2003). *Constructing a language*. Cambridge, MA: Harvard University Press.
Tomasello M & Stahl D (2004). 'Sampling children's spontaneous speech: How much is enough?' *Journal of Child Language 31,* 101–121.
Valian V (1991). 'Syntactic subjects in the early speech of American and Italian children.' *Cognition 40,* 21–81.
Vihman M (1996). *Phonological development*. Oxford: Blackwell.
Weist R, Pawlak A & Carapella J (2004). 'Syntactic-semantic interface in the acquisition of verb morphology.' *Journal of Child Language 31(1),* 31–60.
Wexler K (1998). 'Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage.' *Lingua 106,* 23–79.
Wexler K (2002). 'Lenneberg's dream: learning, normal language development and specific language impairment.' In Schaffer J & Levy Y (eds.) *Language competence across populations: towards a definition of specific language impairment*. Mahwah, NJ: Lawrence Erlbaum.

Language Development: Pre-scientific Studies

R N Campbell, University of Stirling, Stirling, UK

The scientific study of language development may be characterized as involving collection of dated observations linked to general development; the systematic study of such collections by means of type and token counts, tabulation, and so forth; and the development of explanatory accounts grounded in such collections. By the beginning of the 20th century, this sort of apparatus was well established, and works in book form such as Major (1906) and Stern and Stern (1907) presented what are undeniably scientific treatments. However, the foundation of the journal *Pedagogical Seminary* by G. Stanley Hall at Clark University in 1891 provides a more accurate boundary mark: from its foundation onward, the journal featured systematic accounts of aspects of language development.

Early Studies

Reported empirical investigation of language development can probably be dated from the Mughal Akbar the Great's isolation experiment (c. 1578) in which a large number of children were isolated at birth in his palace at Fatehpur Sikri for a period of four years and cared for by unspeaking nurses. According to the three contemporary reports of this experiment, the children did not speak, but may have vocalized. The most explicit account is given in the *Akbarnama*, a life of Akbar prepared by his minister, Sheikh Abu'l Fazl, in the form of an exquisitely illuminated manuscript. Three such manuscripts were produced in the royal workshops. One is located in the Victoria & Albert Museum in London; a second is split between the Chester Beatty Library, Dublin and the British Library; the third is dispersed, with many folios held privately. For printed versions, see Blochmann (1877–86) and Beveridge (1902–1939), who reports the outcome as: "No cry came from that house of silence, nor was any speech heard there. In spite of their four years they had no part of the talisman of speech, and nothing came out except the noise of the dumb." Similar experiments are said to have been carried out at earlier times by the Egyptian Pharaoh Psamtik I in the 7th century B.C., by Emperor Frederick II of Sicily in the 13th century, and by King James IV of Scotland in the late 15th century. However, there are good reasons to doubt the authenticity of all three. For reviews of these early experiments, see Tylor (1878: 79–81), Campbell and Grieve (1982), and Bonvillian *et al.* (1997). It is perhaps remarkable that the first attempt to study language

development should be by means of an experiment (and of course the earlier stories also envisaged experiments), because experimentation appears fairly late in 20th century work. Akbar's experiment may be faulted, because no control group was used. Presumably, it was assumed (perhaps correctly) that a group of newborn children cared for by speaking nurses in a closed house would have spoken in such a manner as to be intelligible to Akbar's court at age four. Akbar's experiment is sufficiently interesting to merit closer examination. There is surely some scope for new insights from the three original versions of the *Akbarnama*, Blochmann's Persian text, and Beveridge's translation.

The 18th Century

The 18th century was a period of strong interest in the nature of language and in its origins. Speculation on these matters abounded among scholars because of the perceived intimate connection with the true nature of Man, and with the distinction between Man and other animals. The best-known discussions are perhaps those of Charles de Brosses (1765), James Burnett [Lord Monboddo](1773–92), Carl Linnaeus, Jean-Jacques Rousseau, and Georges-Louis LeClerc [Comte de Buffon]. Buffon's *Histoire Naturelle* was a huge project that stretched through the second half of the century. Tome 1, of 15 volumes, included his *Natural History of Man and the Animals* and was first published by the Imprimerie Royale in 1749; for accessible English versions, see Barr (1792), Kenrick (1775–76) and Smellie (1781). There is, however, very little in these various works that is grounded in actual observations of children speaking or constrained by other empirical evidence. Buffon's account of infancy (*The Natural History of Man, Section II*) is a notable exception and deserves to be better known. His account is filled with acute observations and shrewd opinions regarding development in the first year, and concludes with a synopsis of early language development. To give the flavor, "At twelve or fifteen months infants begin to lisp. *A* is the vowel which they articulate with most ease, as it requires nothing more than the opening of the lips, and forcing out the breath. *E* requires that the tongue should be raised, as the lips are opened, and is therefore pronounced with a degree or more trouble. *I* is attended with greater difficulty, as in the articulation of this vowel the tongue is elevated still more, and made to approach the teeth of the upper jaw. In the pronunciation of *O* the tongue must be lowered." Buffon proceeds to a least effort theory of the first articulated sounds and words, "in every language *Baba*, *Mama*, *Papa* should be the primitive articulations," anticipating the well-known proposals of Jakobson (1960) by 200 years. A subsequent section uses observations about sound substitution by children to propose a theory of consonant universals and a process of consonant differentiation, once again anticipating some of Jakobson's proposals regarding distinctive features.

A different sort of empirical enquiry was the investigation of the language abilities of apes (then very much a curiosity), and of 'wild' children. The best-known and best-reported cases were those of Marie-Angelique Memmie LeBlanc, found in 1731 near Chalons-sur-Marne, and of Victor, discovered in 1797 near Lacaune (Tarn) and eventually securely held in 1800 near Aveyron. In Mlle. Leblanc's case, the original reports were published in *Mercure de France*, (December 1731, 2983–91), and these suggested that she was speechless and about 10 years old. An account of her life, now thought to have been written by Charles-Marie de La Condamine, appeared as Hecquet (1755), with a later English version by James Burnett (1768), but a more convenient source is Tinland (1970). Mlle. Leblanc appears to have made good progress in learning language and to have been successfully socialized. The story of Victor, and the unsuccessful attempts of Jean-Marc Gaspard Itard to 'civilize' him, is well-known from many recent discussions and versions of Itard's reports. Both cases were the subject of much inconsequential contemporary debate. Tinland (2003) and Douthwaite (2002) provide useful reviews. There is much to be said for the opinion of Bory de Saint-Vincent (1822–31; cited by Tinland): "The discovery (of these savages) first causes a great stir in the newspapers, [but] they end up dying forgotten in some insane asylum. The study of this sort of cripple cannot throw the least light on the primitive state of our species; it is not by means of these exceptions that one must study Man." However, Mlle. LeBlanc is an awkward case for those who believe in a critical period for acquisition.

Finally, Tiedemann (1787) is perhaps the first diary study of development. Versions are available in French and English (Perez, 1881; Tiedemann, 1890; Murchison and Langer, 1927), but the original is hard to locate. It contains a few unsystematic observations relating to language development.

The 19th Century

Interest in developing a scientific approach to psychology gathered pace in the first half of the century, but it was not until the 1870s that scholars in Europe began to make tentative attempts to come to grips with language development. The philosophical journal *Mind* carried short descriptive accounts by Darwin

(1877), Taine (1877), and Pollock (1878). These can be conveniently found, along with Murchison and Langer (1927) and Sully (1881), in Gardner and Gardner (1975). In 1878, the French psychologist Bernard Perez produced the first general text dealing with development and including a long chapter on language development. Perez's text went through numerous revisions, and these were translated into English by Alice M. Christie beginning in 1885. New material was gathered into the language chapter (and old material discarded) between editions. For example, the second edition made considerable use of Egger (1879) and de La Calle (1881), but by the 5th edition of 1890, these sources had been replaced. Perez was shortly followed by the more impressive text of William Preyer in 1882. The second volume of Preyer's text was largely devoted to language development. Like Perez's text, Preyer's was frequently revised and translated, and both stimulated and incorporated new material between revisions, for example, Kussmaul (1874) and the extraordinary investigations of the astronomer Holden (1887). Preyer is unkind to Perez, dismissing him as one who could write that "the glance [is associated with] the magnetic effluvia of the will." However, the two texts were equally influential in stimulating new work and followed similar paths of development. Finally, although James Sully's *Studies in Childhood* lies beyond our boundary of 1891, the language chapters were based on earlier articles by Sully (1881, 1884).

See also: Burnett, James, Monboddo, Lord (1714–1799); Jakobson, Roman (1896–1982); Origin of Language Debate; Psycholinguistics: History; Rousseau, Jean-Jacques (1712–1778).

Bibliography

Barr J S (1792). *Buffon's Natural History.* London: J. S. Barr.

Beveridge H (1902–1939). *Akbarnama* (English version). Bibliotheca Indica Series. Asiatic Society of Bengal. Online text at the Packard Humanities Institute's Persian site http://erga.packhum.org/persian.

Blochmann H (1877–86). *Akbarnama* (Original Persian). Asiatic Society of Bengal.

Bonvillian J D, Garber A M & Dell S B (1997). 'Language origin accounts: was the gesture in the beginning?' *First Language 17,* 219–239.

Burnett J (Lord Monboddo) (1768). *An account of a savage girl, caught wild in the woods of Champagne.* Edinburgh: Kincaid & Bell.

Burnett J (Lord Monboddo) (1773–92). *Of the Origin and Progress of Language* (6 vols). Edinburgh: Kincaid & Creech. Accessible in facsimile reproduction by Scolar Press, 1967. Observations on language development mostly in Volume 1.

Campbell R N & Grieve R (1982). 'Royal investigations of the origin of language.' *Historiographia Linguistica IX,* 43–74.

Darwin C (1877). A Biographical Sketch of an Infant. *Mind 2,* 285–94.

de Brosses C (1765). *Traité de la formation méchanique des langues et des principes physiques de l'étymologie.* Paris: Saillant. Scanned-to-image version at http://gallica.bnf.fr.

de La Calle A (1881). *La Glossologie, essai sur la science expérimentale du langage.* Paris: Maisonneuve.

de Saint-Vincent B (1822–31). *Dictionnaire classique d'histoire naturelle.* Paris: Rey et Gravier.

Douthwaite J V (2002). *The wild girl, natural man, and the monster: dangerous experiments in the Age of Enlightenment.* Chicago: University of Chicago Press.

Egger E (1879). *Observations et réflexions sur le développement de l'intelligence et du langage chez les enfants.* Paris: Picard.

Gardner J K & Gardner H (1975). *First Notes by Observant Parents.* New York: Arno Press.

Hecquet Mme (1755). *Histoire d'une jeune fille sauvage.* Paris. Scanned-to-image version at http://gallica.bnf.fr.

Holden E S (1887). 'On the vocabularies of children under 2.' *Transactions of the American Philological Association 8,* 58–68.

Jakobson R (1960). 'Why 'Mama', why 'Papa'?'. In Kaplan B & Wapner S (eds.) *Perspectives in psychological theory.* International Universities Press. Republished in Jakobson's *Studies on child language and aphasia*, Janua Linguarum Series Minor 114. 21–29. The Hague: Mouton, 1971.

Kenrick W (1775–76). *Buffon's Natural History.* London: T. Bell.

Kussmaul A (1874). Die Storungen der Sprache. In von Ziemssen H W (ed.) *Handbuch der speciellen Pathologie, etc. vol. 14.* Published 1875 in English.

Major D R (1906). *First steps in mental growth; a series of studies in the psychology of infancy.* New York: Macmillan.

Murchison C & Langer S (1927). 'Tiedemann's observations on the mental faculties of children.' *Pedagogical Seminary 34,* 205–230.

Perez B (1878). *Les Trois premières années de l'enfant, étude de psychologie expérimentale.* Paris: Germer Baillière. Later editions to 5th (1892, Paris: Alcan).

Perez B (1881). *Thierri Tiedemann et la science de l'enfant. Mes deux chats, fragment de psychologie comparée.* Paris: Germer Baillière.

Pollock F (1878). An Infant's Progress in Language. *Mind 3,* 392–401.

Preyer W (1882). *Die Seele des Kindes* (2 vols). Leipzig: Graefenhainichen. English version by Brown H W (1890, New York: D. Appleton) reprinted in the Arno Press *Classics in Psychology* series in 1973.

Smellie W (1781). *Buffon's Natural History.* London: Strahan & Cadell.

Stern C & Stern W (1907). *Die Kindersprache.* Leipzig: Barth.

Sully J (1881). Babies in Science. *Cornhill Magazine 43,* 539–554.
Sully J (1884). Baby Linguistics. *English Illustrated Magazine 2,* 110–118.
Taine H (1877). 'On the acquisition of language by children.' *Mind 2,* 252–59. This is a version of an article first published in the *Revue Philosophique* of 1876.
Tiedemann D (1787). *Beobachtungen über die Entwickelung der Seelenfahrigkeiten bei Kindern.* Altenburg: Bonde.
Tiedemann D (1890). *Tiedemann's record of infant life.* English version of Perez (1881) with notes by Soldan F L Syracuse, NY: C.W. Bardeen.
Tinland F (1970). *Histoire d'une jeune fille sauvage.* Bordeaux: Ducros.
Tinland F (2003). *L'Homme sauvage: homo ferus et homo sylvestris, de l'animal à l'homme.* Paris: Editions l'Harmattan (first published 1968, Payot).
Tylor E B (1878). *Researches into the early history of mankind.* London: John Murray.

Relevant Websites

http://erga.packhum.org/persian – Beveridge (1902–39) is available here.
http://gallica,bnf.fr – The original French texts of Tome 1 of Buffon's *Histoire Naturelle* (1749), Hecquet (1755), de Brosses (1765), and the 5th edition (1892) of Perez (1878) are available here in scanned-to-image form.
http://www.jstor.org – Scanned-to-image versions of Darwin (1877), Holden (1877), Taine (1877) and Pollock (1878) are available here.

Language Disorders: 19th Century Studies

P Eling, Radboud University, Nijmegen, The Netherlands

Background

The study of aphasia in the 19th century forms an essential step in our understanding of the various kinds of language pathology. However, the implications for other scientific discussions going on at that time were probably even more important. The Austrian physician and anatomist Franz Joseph Gall (1757–1828) started a discussion on 'localization of mental faculties' around the beginning of the 19th century, which lasted the entire century. Classical (medieval) models localized mental functions (roughly characterized as sensation, imagination (interpretation), and memory) in three cells or ventricles. A popular 18th century view (but dating back to classical Greek views) called 'physiognomy' held that psychological characteristics can be inferred from facial features, the shape of the head in general, and the body. Gall opposed these views and argued that (1) the material substance of the brain, in particular, the cortex, forms the basis of mental functions (instead of the cavities inside), and that (2) each mental faculty has its own seat, a circumscribed area on the cortex. Gall's view of the mind, craniology, later phrenology, seems inspired by the faculty psychology of Thomas Reid rather than the classical ventricular classification of faculties or the physiognomic framework of characters. These were revolutionary claims, both for the religious and secular authorities, but also for investigators of the nervous system. Gall's views became very popular in high-society circles, and eventually the middle class as well, due particularly to the efforts of Gall's assistant, Johann Caspar Spurzheim (1776–1832), and the various phrenological societies. This lasted until 1861, when the principle of localization of function became acceptable in scientific circles. The study of the effects of lesions to various parts of the brain on language behavior produced important empirical evidence for and against this localization principle. Aphasiologists were highly aware of this additional feature of their work.

A related but separate issue forms the development of 'connectionistic models.' Before the 19th century, mental functions had been studied mostly by philosophers. It had been argued by Descartes and De la Mettrie in the 17th century that the brain could be regarded as a machine. This mechanistic approach led to the development of 'process models' of language production and comprehension in the second half of the 19th century. These models distinguished various centers representing sounds, words, or concepts, for instance, and 'pathways' of nerve fibers. A process model thus described an author's view on how a word was transformed into a (meaningful) linguistic representation and subsequently into speech or writing. Once this approach of analyzing language disorders in relation to brain lesions became fashionable, aphasiologists discussed the nature of the units and the various connections. A center was connected to one or more others, and a lesion in either a center or a pathway would result in the loss of a particular function or of the transmitting of outputs of a center.

Sully J (1881). Babies in Science. *Cornhill Magazine 43,* 539–554.
Sully J (1884). Baby Linguistics. *English Illustrated Magazine 2,* 110–118.
Taine H (1877). 'On the acquisition of language by children.' *Mind 2,* 252–59. This is a version of an article first published in the *Revue Philosophique* of 1876.
Tiedemann D (1787). *Beobachtungen über die Entwickelung der Seelenfahrigkeiten bei Kindern.* Altenburg: Bonde.
Tiedemann D (1890). *Tiedemann's record of infant life.* English version of Perez (1881) with notes by Soldan F L Syracuse, NY: C.W. Bardeen.
Tinland F (1970). *Histoire d'une jeune fille sauvage.* Bordeaux: Ducros.
Tinland F (2003). *L'Homme sauvage: homo ferus et homo sylvestris, de l'animal à l'homme.* Paris: Editions l'Harmattan (first published 1968, Payot).
Tylor E B (1878). *Researches into the early history of mankind.* London: John Murray.

Relevant Websites

http://erga.packhum.org/persian – Beveridge (1902–39) is available here.
http://gallica,bnf.fr – The original French texts of Tome 1 of Buffon's *Histoire Naturelle* (1749), Hecquet (1755), de Brosses (1765), and the 5th edition (1892) of Perez (1878) are available here in scanned-to-image form.
http://www.jstor.org – Scanned-to-image versions of Darwin (1877), Holden (1877), Taine (1877) and Pollock (1878) are available here.

Language Disorders: 19th Century Studies

P Eling, Radboud University, Nijmegen, The Netherlands

Background

The study of aphasia in the 19th century forms an essential step in our understanding of the various kinds of language pathology. However, the implications for other scientific discussions going on at that time were probably even more important. The Austrian physician and anatomist Franz Joseph Gall (1757–1828) started a discussion on 'localization of mental faculties' around the beginning of the 19th century, which lasted the entire century. Classical (medieval) models localized mental functions (roughly characterized as sensation, imagination (interpretation), and memory) in three cells or ventricles. A popular 18th century view (but dating back to classical Greek views) called 'physiognomy' held that psychological characteristics can be inferred from facial features, the shape of the head in general, and the body. Gall opposed these views and argued that (1) the material substance of the brain, in particular, the cortex, forms the basis of mental functions (instead of the cavities inside), and that (2) each mental faculty has its own seat, a circumscribed area on the cortex. Gall's view of the mind, craniology, later phrenology, seems inspired by the faculty psychology of Thomas Reid rather than the classical ventricular classification of faculties or the physiognomic framework of characters. These were revolutionary claims, both for the religious and secular authorities, but also for investigators of the nervous system. Gall's views became very popular in high-society circles, and eventually the middle class as well, due particularly to the efforts of Gall's assistant, Johann Caspar Spurzheim (1776–1832), and the various phrenological societies. This lasted until 1861, when the principle of localization of function became acceptable in scientific circles. The study of the effects of lesions to various parts of the brain on language behavior produced important empirical evidence for and against this localization principle. Aphasiologists were highly aware of this additional feature of their work.

A related but separate issue forms the development of 'connectionistic models.' Before the 19th century, mental functions had been studied mostly by philosophers. It had been argued by Descartes and De la Mettrie in the 17th century that the brain could be regarded as a machine. This mechanistic approach led to the development of 'process models' of language production and comprehension in the second half of the 19th century. These models distinguished various centers representing sounds, words, or concepts, for instance, and 'pathways' of nerve fibers. A process model thus described an author's view on how a word was transformed into a (meaningful) linguistic representation and subsequently into speech or writing. Once this approach of analyzing language disorders in relation to brain lesions became fashionable, aphasiologists discussed the nature of the units and the various connections. A center was connected to one or more others, and a lesion in either a center or a pathway would result in the loss of a particular function or of the transmitting of outputs of a center.

The notion of 'connectionistic models' derives from this general way of thinking and is highly confusing to those familiar with current dynamic systems approaches, which are also sometimes referred to as 'connectionistic' (Caplan, 1987).

Localization of Speech

Many articles on language pathology up to 1865 discussed the validity of the localization principle. Jean Baptiste Bouillaud (1796–1881) had collected a large set of data on patients with speech problems. He claimed that in the majority of cases the lesion was localized in the 'anterior' part of the brain. This was in agreement with Gall's system: Gall localized the language faculty, verbal memory, right behind the eye sockets. At that time, the distinction between anterior and posterior was as precise as one could be; up to that time, the cortex had been considered irrelevant and no further description and division was required of the seemingly randomly patterned gyri and sulci. Bouillaud's repeated attempts from the 1820s to the 1850s remained unconvincing to the medical and scientific establishment.

A discussion on localization flared up in the newly founded Anthropological Society in Paris, whose members were mostly physicians. Its secretary was Pierre Paul Broca (1824–1880). In 1861 he claimed that one of his patients, who had suffered from a severe speech problem (all he could utter was the syllable 'Tan') had a lesion in the anterior part of the brain. Twenty days later, Tan died and Broca showed his brain to his colleagues, pointing out that a lesion in the anterior part of the brain leads to 'aphemia,' a disturbance in the articulation of words. Trousseau suggested using the term 'aphasia' in 1864. The case of Tan apparently was convincing; the notion that mental faculties are somehow seated in the cortex was quickly accepted by many physicians. Yet John Hughlings-Jackson (1835–1911) warned that to localize a symptom (e.g., aphasia) is not the same as to localize a function. Being educated in a completely different philosophical tradition, associationism, he could not accept the idea that a language function, in Broca's case, the articulation of speech, could be localized in a circumscribed cortical area.

Localization of Language Comprehension

One of the first clear descriptions of language comprehension problems due to a brain disorder can be found in Jacques Lordat's (1773–1870) description in 1843 of his own aphasia. Until then, language comprehension problems were usually ascribed to the deterioration of general intellect. The French physician Francois Baillarger (1809–1890) described in 1865 a woman with aphasia who had been regarded as deaf and insane, but it was soon demonstrated that neither deafness nor insanity could explain the comprehension problems. Heavily influenced by the Austrian neuroanatomist Theodor Meynert (1833–1892), Carl Wernicke (1848–1905) wrote a dissertation in 1874 claiming that speech perception is localized in a posterior part of the brain, the temporal lobe, in an area now known as Wernicke's area. Moreover, he claimed that the site for language could be conceived of as consisting of two centers that were connected to each other – one for perception and one for speech production. Lesions in either one of these centers or in the connecting pathways would result in different patterns of language pathology.

Diagrams

To illustrate his theory, Wernicke produced a drawing of a cerebral hemisphere with two circles that represented the centers for word perception and word production. The centers had connections to each other and to the cranial nerves in the brain stem. A similar 'diagram' had been proposed by Adolf Baginsky (1843–1915) in 1871 in a somewhat neglected paper. A crucial difference between these two theories was Baginsky's assumption that there was a center for concepts independent from that for (auditory or visual) word forms; Wernicke explicitly denied this and argued that the brain registers only sound and visual images. Wernicke's scheme was elaborated by Ludwig Lichtheim (1885) and is still being used today. In England, Henry Charlton Bastian had developed a similar diagram, and in France, Jean Martin Charcot developed the famous 'clock diagram.'

The general idea of representing language as a set of interconnected circles became very popular and produced lots of 'diagram makers,' as Henry Head called them in his famous monograph. Essentially, this discussion only partly dealt with **where** specific mental faculties are represented. Rather, the nature of a 'function' itself was at stake. For instance, is it necessary to presuppose that apart from a center for words (that is, for word forms), there is also a center for meaning? Adolf Kussmaul's discussion of these models in 1877 is an excellent example of how these process models could explain aphasic behavior in a functional way.

Localization of Reading and Writing

Assuming the existence of various language centers, William Ogle was one of the first to argue that apart from speech problems, brain lesions can result in

distinctive problems in writing: agraphia. In one form, patients can write, but they make errors – the order of letters is disturbed, or nonexisting words are produced. In another form, writing is no longer possible. Aphasiologists working with diagrams also examined the reading of isolated words and often noted reading problems in their patients. They often assumed a separate center for the recognition of visual words – that is, for reading. Joseph Jules Dejerine (1849–1917) contributed to our understanding of reading and writing disorders in the presence of brain lesions. In 1887, Dejerine distinguished four language centers, including a center for reading (localized in the angular gyrus) and one for writing (localized at the foot of the second frontal gyrus). He later subscribed to the claims of Wernicke and Lichtheim that deficits in writing are due to problems with 'inner speech.' In 1882, Dejerine argued that there are two forms of reading disorders: one accompanied by writing deficits and another that is a pure form of word blindness in which the patient indeed can write correctly but cannot read what he has written.

Cerebral Dominance

The discussion on localization of function concentrated on specific parts of the cerebral hemispheres that presumably subserved a specific mental function. It took several decades to convince the scientific community of the validity of this principle, but once this step was taken, it was evidently no problem to accept another principle of localization, namely, that of hemisphere asymmetry. After Broca had shown the brain of Tan to the Anthropological Society, he began to collect additional cases. These not only seemed to confirm his claims, but also pointed to a rather surprising conclusion. It appeared that only the left hemisphere was involved in speech production, at least in most individuals. From a physiological point of view, this was hard to believe: form and function go together and since the two hemispheres look similar, they should serve the same functions, like other paired organs. Yet when Broca introduced the principle of hemisphere asymmetry in 1865, it was little disputed. On the contrary, it was immediately related to hand preference, and the two phenomena were regarded as signs of 'cerebral dominance,' the notion that one hemisphere – usually the left – develops quicker and therefore becomes the dominant one.

A Linguistic Approach

The physicians quarreled among themselves about the number and nature of the language centers, but they practically always assumed that the basic language unit was the word, a noun or a verb. Ferdinand Finkelnburg was an exception, claiming in 1870 that the core problem is an 'asymbolia,' the incapacity to deal with symbols of any kind. Nobody worried about grammar until Heymann Steinthal (1871) decided to apply his views on linguistic theory to language disorders. Linguistics at that time was basically comparative linguistics: looking for the derivation of current words from older languages. Steinthal pleaded for a different approach, one we would refer to as 'psycholinguistics.' He argued that language is a means to express an idea, usually in the form of a sentence. The idea or message is 'unpacked' by psychological processes, delivering a linguistic structure, which calls for words, words with specific focus (morphology), and these are then produced as sounds. Now one could assume a number of disorders at different stages in this process of 'unpacking' resulting in disorders that are not modality-specific (input–output, auditory–visual) but are related to levels of representation: message, sentence, word, sound level. When Steinthal and his friend, the psychologist Lazarus, finally got their audience with the physicians in the Berlin academic hospital, the latter indicated that the psychologists had no idea what aphasia actually means, and they should better start visiting the clinics to become acquainted with it. Unfortunately, Steinthal's psycholinguistic approach was never taken seriously until late in the 20th century.

John Hughlings Jackson also claimed that the basic unit is not a simple word but 'the proposition.' A significant observation, but one not transformed in a testable view; at least, few physicians picked up the idea and applied it to their studies of language pathology. Unfortunately, 19th century aphasiologists remained much more interested in issues of localization than in the psycholinguistic principles underlying language behavior and language disorders.

See also: Aphasia Syndromes; Associationism and Connectionism.

Bibliography

Boller F (1978). 'Comprehension disorders in aphasia: a historical review.' *Brain and Language 5,* 149–165.

Caplan D (1987). *Neurolinguistics and linguistic aphasiology. An introduction.* Cambridge: Cambridge University Press.

Clarke E & Jacyna L S (1987). *Nineteenth-century origins of neuroscientific concepts.* Berkeley: University of California Press.

De Bleser R (1987). 'From agrammatism to paragrammatism: German aphasiological traditions and grammatical disturbances.' *Cognitive Neuropsychology 4,* 187–256.

De Bleser R, Cubelli R & Luzzalli C (1993). 'Conduction aphasia, misrepresentations and word representations.' *Brain and Language 45,* 475–494.
De Bleser R & Luzzalli C (1989). 'Models of reading and writing and their disorders in classical German aphasiology.' *Cognitive Neuropsychology 6,* 501–513.
Eling P (ed.) (1994). *Reader in the history of aphasia.* Amsterdam: John Benjamins.
Finger S (1994). *Origins of neuroscience. A history of explorations into brain function.* Oxford: Oxford University Press.
Kolle K (ed.) (1970). *Grosse Nervenärzte* (3 vols). Stuttgart: Thieme.
Kussmaul A (1877). *Die Störungen der Sprache. Versuch einer Pathologie der Sprache.* Leipzig: Vogl.
Lesky E (1979). *Franz Joseph Gall. Naturforscher und Anthropologe.* Bern: Huber.
Tesak J (2001). *Geschichte der Aphasie.* Idstein: Schulz-Kirchner Verlag.
Steinthal H (1871). *Einleitung in die Psychologie und Sprachwissenschaft.* Berlin: Dümmler's Verlagsbuchhandlung.
Whitaker H A (1988). 'Neurolinguistics from the middle ages to the pre-modern era: historical vignettes.' In Stemmer B & Whitaker H A (eds.) *Handbook of neurolinguistics.* San Diego: Academic Press. 27–54.

Language Disorders: 20th-Century Studies, Pre-1980

P Eling, University of Radboud, Nijmegen, The Netherlands

From Localization to a Holistic Approach

The 'classical' approach, stimulated by the works of Wernicke (1874) and Lichtheim (1885), aimed at the localization of specific language centers in cortical areas: modality-dependent word input and output centers, without (Wernicke) or with (Lichtheim) a special center for (abstract) concepts. Some of Wernicke's pupils adhered to this model. Karl Kleist (1870–1962) summarized his studies of a large number of soldiers with penetrating brain wounds in his *Gehirnpathologie* in 1934. He believed that it was possible to make a precise localization of the cortical regions underlying diverse language functions. Probably the most extreme attempts to localize language function were performed by Samuel Henschen (1847–1930). In France, Jules Joseph Dejerine (1849–1917) was the most influential localizationist in the beginning of the 20th century.

At that time, opposition arose towards a simplistic, mechanistic view of language based on associations between sensory and motor representations of words. Hughlings Jackson (1835–1911) at first had been enthusiastic about Broca's claims concerning the localization of the mechanism to articulate speech. However, he soon realized that Broca's views on language differed from his own. For instance, Jackson could not accept the notion of circumscribed centers. Moreover, he believed the proposition to be the core linguistic unit that was affected, not individual words.

Pierre Marie (1853–1940), originally educated in the diagram tradition, started the 'revolution' in 1906 in France by claiming that Broca's area, the third frontal gyrus, plays no role whatsoever in language. The claim was based on contradictory evidence: patients with lesions in Broca's area and no aphasia, and patients with Broca's aphasia due to a lesion in another part of the brain. Marie argued that aphasia is the result of a lesion in the temporal lobe producing an intellectual impairment, a reduction of all knowledge and competence acquired didactically; there is only one form of aphasia, which can vary in severity. Broca's aphasia is merely aphasia with an added articulatory disorder. In 1908, the 'clash,' an extensive debate between the localizationists (headed by Dejerine), and the antilocalizationists (headed by Marie), ended in a draw.

The Swiss neurologist Constantin von Monakow (1853–1930) also rejected the notion of centers in the brain. He used the term 'language fields' instead. He also regarded the area surrounding the Sylvian fissure as important for language, but there are no strict borders: more distant areas can interact with these language areas. He convinced many of the inadequacy of a strict localization view by comparing the localization of language issue to the question of where the melody is localized in a music box.

Henry Head (1891–1940) was a great admirer of Jackson; in his book titled *Aphasia and kindred disorders of speech*, published in 1926, he severely criticized the 'diagram makers,' a nickname that proved to be very effective in disqualifying the entire localizationist approach. Like Marie, Head believed that essentially aphasia is a disorder in the formation and expression of linguistic and nonlinguistic symbols. He explicitly extended the disorder to other cognitive deficits such as problems in arithmetic. Head clinically distinguished four types of aphasic disorders: at the verbal level (word form), nominal level (meaning of

De Bleser R, Cubelli R & Luzzalli C (1993). 'Conduction aphasia, misrepresentations and word representations.' *Brain and Language 45,* 475–494.
De Bleser R & Luzzalli C (1989). 'Models of reading and writing and their disorders in classical German aphasiology.' *Cognitive Neuropsychology 6,* 501–513.
Eling P (ed.) (1994). *Reader in the history of aphasia.* Amsterdam: John Benjamins.
Finger S (1994). *Origins of neuroscience. A history of explorations into brain function.* Oxford: Oxford University Press.
Kolle K (ed.) (1970). *Grosse Nervenärzte* (3 vols). Stuttgart: Thieme.
Kussmaul A (1877). *Die Störungen der Sprache. Versuch einer Pathologie der Sprache.* Leipzig: Vogl.
Lesky E (1979). *Franz Joseph Gall. Naturforscher und Anthropologe.* Bern: Huber.
Tesak J (2001). *Geschichte der Aphasie.* Idstein: Schulz-Kirchner Verlag.
Steinthal H (1871). *Einleitung in die Psychologie und Sprachwissenschaft.* Berlin: Dümmler's Verlagsbuchhandlung.
Whitaker H A (1988). 'Neurolinguistics from the middle ages to the pre-modern era: historical vignettes.' In Stemmer B & Whitaker H A (eds.) *Handbook of neurolinguistics.* San Diego: Academic Press. 27–54.

Language Disorders: 20th-Century Studies, Pre-1980

P Eling, University of Radboud, Nijmegen, The Netherlands

From Localization to a Holistic Approach

The 'classical' approach, stimulated by the works of Wernicke (1874) and Lichtheim (1885), aimed at the localization of specific language centers in cortical areas: modality-dependent word input and output centers, without (Wernicke) or with (Lichtheim) a special center for (abstract) concepts. Some of Wernicke's pupils adhered to this model. Karl Kleist (1870–1962) summarized his studies of a large number of soldiers with penetrating brain wounds in his *Gehirnpathologie* in 1934. He believed that it was possible to make a precise localization of the cortical regions underlying diverse language functions. Probably the most extreme attempts to localize language function were performed by Samuel Henschen (1847–1930). In France, Jules Joseph Dejerine (1849–1917) was the most influential localizationist in the beginning of the 20th century.

At that time, opposition arose towards a simplistic, mechanistic view of language based on associations between sensory and motor representations of words. Hughlings Jackson (1835–1911) at first had been enthusiastic about Broca's claims concerning the localization of the mechanism to articulate speech. However, he soon realized that Broca's views on language differed from his own. For instance, Jackson could not accept the notion of circumscribed centers. Moreover, he believed the proposition to be the core linguistic unit that was affected, not individual words.

Pierre Marie (1853–1940), originally educated in the diagram tradition, started the 'revolution' in 1906 in France by claiming that Broca's area, the third frontal gyrus, plays no role whatsoever in language. The claim was based on contradictory evidence: patients with lesions in Broca's area and no aphasia, and patients with Broca's aphasia due to a lesion in another part of the brain. Marie argued that aphasia is the result of a lesion in the temporal lobe producing an intellectual impairment, a reduction of all knowledge and competence acquired didactically; there is only one form of aphasia, which can vary in severity. Broca's aphasia is merely aphasia with an added articulatory disorder. In 1908, the 'clash,' an extensive debate between the localizationists (headed by Dejerine), and the antilocalizationists (headed by Marie), ended in a draw.

The Swiss neurologist Constantin von Monakow (1853–1930) also rejected the notion of centers in the brain. He used the term 'language fields' instead. He also regarded the area surrounding the Sylvian fissure as important for language, but there are no strict borders: more distant areas can interact with these language areas. He convinced many of the inadequacy of a strict localization view by comparing the localization of language issue to the question of where the melody is localized in a music box.

Henry Head (1891–1940) was a great admirer of Jackson; in his book titled *Aphasia and kindred disorders of speech*, published in 1926, he severely criticized the 'diagram makers,' a nickname that proved to be very effective in disqualifying the entire localizationist approach. Like Marie, Head believed that essentially aphasia is a disorder in the formation and expression of linguistic and nonlinguistic symbols. He explicitly extended the disorder to other cognitive deficits such as problems in arithmetic. Head clinically distinguished four types of aphasic disorders: at the verbal level (word form), nominal level (meaning of

word or sentence), syntactic level, and semantic problems, in particular nonliteral meanings of words or expressions and of intentions.

Adherents of a holistic approach not only rejected the notion of circumscribed centers, but they also stressed that language was much more than words associated to images. Meaning, in the most extended way (connecting language to attitudes, feelings, intentions, and motives), was more fundamental to them. The most famous 'holist' was probably the German neuropsychiatrist Kurt Goldstein (1989–1965). One of the last pupils of Wernicke, he became director of the Institute for Research on the After-Effects of Brain Injury in Frankfurt, a rehabilitation center for soldiers of World War I. He moved to the United States in 1935. Goldstein rejected Wernicke's view of acoustic and motor images as the basis of words as not sufficiently psychological. Language is about meaning. We are conscious of meanings, not of acoustic images, he argued. Interestingly, Goldstein, like Head, based his views on Jackson's. According to Goldstein, the loss of the 'abstract attitude' is the essential deficit in aphasia. In the 'concrete attitude,' environmental stimulation determines our thinking and actions. However, when we abstain from environmentally determined behavior, we act according to an abstract attitude. In this sense, the abstract attitude may be compared to Head's symbolic behavior; it is much wider than language in the usual meaning of the word.

Grammar and Aphasia

Within the associationistic (Wundtian) approach to language and the brain, the word was the central unit, at least for most aphasiologists. In the 'Würzburger School' (O. Külpe, K. Bühler), a different view of language processing was developed in which the whole sentence became the central unit. At the beginning of the 20th century, aphasiologists also noted that some patients had problems, not just with speech production or perception, but with speaking correctly, a phenomenon labeled 'agrammatism.'

Arnold Pick (1851–1924) claimed that agrammatism arises due to a lesion in the left temporal lobe that affects speech production processes rather than higher cognitive processes. Karl Bonhoeffer (1868–1948) regarded agrammatic speech as a consequence of a lesion in Broca's speech center that resulted in laborious speech. Karl Heilbronner (1869–1914), on the other hand, argued that agrammatism is an independent phenomenon. In addition, he claimed that severe agrammatic speech can occur in patients who do not show any deficit in comprehension and that agrammatism is a grammar problem that is only present language production, not perception. In his famous chapter, *Die agrammatischen Sprachstörungen*' (1913), Pick formulated an elaborate model of sentence production, distinguishing between a syntactic representation (*Satzschema*, sentence scheme) and word forms that are inserted in the sentence scheme. Pick noticed that in agrammatic speech the phrases that are produced are essentially intact. Therefore, what is produced is not the normal sentence but an adapted version, *Notsprache* (emergency speech), which requires less effort. Sentences are shorter and syntactically less specified. Max Isserlin (1879–1941) defended a similar position.

Kleist in 1914 introduced the notion of 'paragrammatism,' which he regarded as a different type of grammatical problem, one of word order. In his view, agrammatism is a kind of simplification of the word order in which subordinate clauses do not occur. In paragrammatism, the capacity to produce certain syntactic constructions is not lost, but the production of phrases is often interrupted or may be combined with alternative expressions.

This discussion on agrammatism was started again in the 1970s with arguments that were not too different: modality specificity, economy of speech, and deficit vs. adaptation.

Clinical Approach

Single case studies usually formed the basis of research by most aphasiologists in the first part of the 20th century. The American neurologist Theodore Weisenburg (1875–1934) and the psychologist Katharine McBride tried to solve the question of dissociation of aphasic phenomena in a 'clinical way.' They developed a large test battery and published data from large numbers of patients in their work, *Aphasia: a clinical and psychological study* (1935). To establish which errors or what level of errors is normal, they also examined healthy controls. The classification of aphasic syndromes 'emerging' from the results was as follows: primarily motoric, primarily sensory, sensory-motor, and amnestic aphasia (word-finding problems in speech production). This shift from the single case to group studies announced another debate that also dominated the neurolinguistic discussion in the final quarter of the century. Joseph Wepman (b. 1907), professor in psychology, and Hildred Schuell (b. 1907), a clinical speech pathologist, worked independently of each other in this clinical tradition and developed rehabilitation programs for aphasic patients. They also favored the holistic approach rather than the speech centers approach. Wepman developed the Language Modalities Test for Aphasia (1959), while Schuell

introduced the Minnesota Test for the Differential Diagnosis of Aphasia (1966). An important and somewhat surprising finding coming out of these group studies was that language representation in left-handers is not a mirror image of that in right-handers.

From Aphasia to Neurolinguistics

The American neurologist Norman Geschwind (1926–1984) read much of Wernicke's works during his medical training, which helped stimulate him to found the Boston University Aphasia Research Center in the early 1960s. His plea for the reintroduction of the 'disconnection approach' as a means for the study of the relation between language (functions) and the brain (centers) had a strong impact on the development of the interest in neuropsychology, in general, and aphasia, in particular. Many famous neurolinguists came to Boston for training under the supervision of Harold Goodglass (1920–2002). This renewed interest in the neuropsychology of language soon led to a shift of focus from the word (in Geschwind's approach) to the sentence level. Goodglass, and to a larger extent Alfonso Caramazza and Edgar Zurif, triggered this movement in the 1970s when they tried to translate the insights of the American linguist Noam Chomsky (b. 1928), who had dominated the linguistics scene since 1965 with his universal transformational grammar theory, to aphasic language performance. In his dissertation (1969), Harry Whitaker explicitly applied transformational grammar to the study of aphasia cases. Aphasia research finally became the interdisciplinary study it should be, combining linguistic (representation), psychological (processes), and neurological (anatomical and physiological mechanisms) models. This is expressed in the new label for aphasiology: neurolinguistics, a term widely used by the Russian neuropsychologist Alexander Luria and the French neurologist Henri Hécaen in the 1960s.

Roman Jacobson (1896–1982) had already introduced a neurolinguistic analysis of aphasic speech in the 1940s. He was Russian in origin, but from 1941 he worked as a linguist in America. He was interested in the sound structure (phonology) of languages, particularly Russian. He analyzed its evolution in children's speech and compared the changes he observed there to speech errors in aphasia. He argued that there are two fundamental dimensions along which aphasic speech can be classified: disorders of 'similarity' (replacements or selection errors) and disorders of 'contiguity,' temporal ordering problems. His work had little impact originally, but was later rediscovered and applied primarily to phonological analysis (such as that of his student Sheila Blumstein) of speech errors.

Epilepsy and Language

Neurosurgery developed significantly in the 20th century and offered important new options for treatment of patients with severe forms of epilepsy, including removal of brain tissue in which the epileptic focus was localized. To prevent interference with language processes, neurosurgeons ascertained whether the affected areas were involved in language. To that end, they applied electrical stimulation to the exposed cortex and asked patients to perform a short series of language tasks (e.g., naming, counting, sentence-picture matching). The American neurologist Wilder Penfield (1891–1976), who was trained in Germany by Otfried Foerster and worked in Montreal, Canada, described (with his collaborator, Lamar Roberts) the results of their long-lasting research in a monograph, *Speech mechanisms and the brain*, (1959). Among other things, they noted that not just the classical (Broca and Wernicke) areas, but a wider area appeared associated with language functions, including the supplementary motor area. Stimulation sometimes produced positive effects (e.g., vocalizations), and sometimes it interfered with ongoing language processes. George Ojemann followed up on this line of research and showed that stimulation of subcortical areas also affected language. The Penfield and Roberts data also indicated that language was localized in the left hemisphere in approximately 70% of the left-handers. This puzzling relationship between handedness and side of language representation was also found in studies using the WADA-technique, a procedure introduced by Juhn Wada (b. 1924) in 1949 in Japan and later applied as a diagnostic procedure in epilepsy neurosurgery when Wada collaborated with Penfield in Montreal.

Neurolinguistic Publications

The renewed interest in aphasiology in the 1960s attracted the attention of students of various disciplines. The tremendous scientific activity in this area is reflected in the sharp increase in numbers of papers, books, journals, conferences, and societies devoted to neurolinguistics. For example, the major (annual) conference devoted to aphasia, the Academy of Apahasia, was founded in 1962. Another conference, BABBLE (Body for the Advancement of Brain Behavior and Language Enterprises), started in 1976. The journal *Brain and Language* was founded by Harry Whitaker in 1974, and he was also the co-editor (with H. A. Whitaker) of a book series, *Studies in*

neurolinguistics (1976–1979). Another series of edited books on neurolinguistics started in 1973. As in the 19th century, rapid developments in the area of aphasia triggered similar developments in the study of other cognitive disorders following brain disorders.

Bibliography

Arbib M, Caplan D & Marshall J C (1982). 'Neurolinguistics in Historical Perspective.' In Arbib M, Caplan D & Marshall J C (eds.) *Neural models of language process.* San Diego: Academic Press.

Bleser R de (1987). 'From agrammatism to paragrammatism: German aphasiological traditions and grammatical disturbances.' *Cognitive Neuropsychology 4,* 187–256.

Caplan D (1994). 'Language and the brain.' In Gernsbacher M A (ed.) *Handbook of psycholinguistics.* San Diego: Academic Press.

Caplan D (1987). *Neurolinguistics and linguistic aphasiology: an introduction.* New York: Cambridge University Press.

Eling P (ed.) (1994). *Reader in the history of aphasia.* Amsterdam: Benjamins.

Finger S (1994). *Origins of neuroscience: a history of explorations into brain function.* Oxford: Oxford University Press.

Glozman J (1996). 'A brief history of Russian aphasiology.' *Neuropsychology Review 6,* 1–10.

Goodglass H & Wingfield A (1988). 'The changing relationship between anatomic and cognitive explanation in the neuropsychology of language.' *Journal of Psycholinguistic Research 27,* 147–165.

Goodglass H (1988). Historical perspective on concepts of aphasia. In Boller F & Grafman J (eds.) *Handbook of neuropsychology,* vol. 1. Amsterdam: Elsevier.

Lecours A R, Chain F, Poucet M, Nespoulous J L & Joanett Y (1992). 'Paris 1908: The hot summer of aphasiology or a season in the life of a chair.' *Brain and Language 42,* 105–152.

Lichtheim H (1885). 'On aphasia.' *Brain 7,* 433–485.

Noppeney U & Wallesch (2000). 'Language and cognition. Kurt Goldstein's theory of semantics.' *Brain and Cognition 44,* 367–386.

Penfield W & Roberts L (1959). *Speech and brain mechanisms.* New Jersey: Princeton University.

Tesak J (2001). *Geschichte der aphasie.* Idstein: Schulz-Kirchner Verlag.

Weisenburg T & McBride K (1935). *Aphasia: a clinical and psychological study.* New York: The Commonwealth Fund.

Wernicke C (1874). *Der Aphasichse Symptomencomplex. Eine psychologische Studie auf anatomischer Basis.* Breslau: Cohn & Weigert.

Whitaker H & Whitaker H A (1976–1979). *Studies in Neurolinguistics* (Vols 1–4). New York: Academic Press.

Whitaker H A (1988). 'Neurolinguistics from the middle ages to the premodern era: historical vignettes.' In Stemmer B & Whitaker H A (eds.) *Handbook of neurolinguistics.* San Diego: Academic Press. 27–54.

Zurif E & Swinney D (1994). 'The neuropsychology of language.' In Gernsbacher M A (ed.) *Handbook of psycholinguistics.* San Diego: Academic Press.

Language Education and Gender

B Egbo, University of Windsor, Windsor, Ontario, Canada

Research in language and gender as a subfield in sociolinguistics has accelerated during the last several decades largely due to the critical work and writings of feminist sociologists, anthropologists, sociolinguists, and post-structuralists (e.g., Lakoff, 1975; Philips *et al*., 1987; Coates, 1993), much in the same way as Simone de Beauvoir's seminal work *The second sex*, released in 1953, galvanized contemporary feminist movement. Emerging naturally from the interest in language and gender have been inquiries that collectively cohere around a critical analysis of the interface between language, education, and gender, which is the subject of this review.

Different from its classificatory or grammatical meaning, gender is examined here in terms of a social construct that society arbitrarily ascribes to men and women, based on their perceived roles in society. Such a framework recognizes that while the biological state of 'femaleness' and 'maleness' is, recent claims to the contrary notwithstanding (see Bergvall *et al*., 1996), stable and universal, understandings of gender are in a constant state of flux, temporal, depend on the prevailing social order, and mediate the construction of the gendered self, which in turn, shapes individual identity (Bjerrum Nielsen and Davies, 1997; Egbo, 2004).

While it is not clear how it happens, given the polemic among sex-role development theorists, researchers agree that cross-culturally, by the time children begin their formal schooling, they have developed an emotional attachment to their gender, and already act, speak, and behave according to conventional images of gender – behaviors and discourse norms that are further reified in schools (Spender, 1982; Bjerrum Nielsen and Davies, 1997).

Passy P (1899). *De la methode directe dans l'enseignement des langues vivante*. Paris: Colin.
Richards J C & Nunan D (eds.) (1990). *Second language teacher education*. Cambridge: Cambridge University Press.
Rivers W (1968). *Teaching foreign language skills*. Chicago: Chicago University Press.
Stevick E (1976). *Memory, meaning and method*. Rowley Mass: Newbury House.
Thorndike E L (1932). *A teacher's word book of 20,000 words*. New York: Teachers' College, Columbia University.
Trim J L M, Richterich R, van Ek J A & Wilkins D A (1973). *Systems development in adult language learning*. Strasbourg: The Council of Europe. (Reprinted 1980, Oxford: Pergamon Press.)
West M (1953). *A general service list of English words*. London: Longman.

Language Education – Vocabulary

P Nation, Victoria University of Wellington, New Zealand

What Vocabulary?

There are numerous vocabulary counting programs that change texts into frequency lists of vocabulary and show how much text is covered by frequency-ranked lists. What these counts show is that a relatively small number of words account for a very large proportion of text. The 270 function word types (176 word families), such as *one, a, the, because, in, must*, cover about 44% of the running words (tokens) found in most texts. These function words, however, make up a small proportion of the frequent words of English. Among the highly ranked items are words such as *time*, *say*, *day*. The most frequent 2000 words of English, which include most of the function words, are the essential widely used words of the language. They are important no matter what use is made of the language, and they cover a large proportion of spoken and written text. They also represent a feasible learning goal for an English course of 800 to 1000 hours.

For learners with academic purposes such as study through the medium of English in senior high school or university, the next important vocabulary learning goal is academic vocabulary. The best researched list of these words is the Academic Word List (Coxhead, 2000), which consists of 570 word families arranged into ten sublists with sublist 1 containing the most frequent widest-range items. There are well-researched tests available (Schmitt *et al.*, 2001; Nation, 2001) called the Vocabulary Levels Tests, which can be used to quickly determine which level of vocabulary learners need to focus on. The main distinction to be made when interpreting the results of the test is between the high frequency words (the first 2000 words and the Academic Word List) and the low frequency words (3000 level onward).

How Should Words Be Dealt With?

The high frequency/low frequency distinction is an important one when planning a vocabulary program. This is because from a teaching perspective these two groups of words should be dealt with in different ways. The 2000 or 2570 high frequency words deserve classroom time. This is because of their high frequency and wide occurrence. This classroom time can involve direct teaching of these words, doing activities designed to teach and practice these words, and deliberate learning of these words. Low frequency words on the other hand do not deserve classroom time. This is because there are so many of them and because of their low frequency of occurrence. When learners know the high frequency words, they need to begin learning low frequency words. However, the teacher should not be spending valuable classroom time teaching those low frequency words. Instead, the teacher should spend time teaching the four most useful strategies for dealing with these words, namely the strategies of guessing words from context clues, deliberate learning using word cards, helping memory by using word part analysis, and dictionary use.

The opportunities for teaching and learning high frequency vocabulary can be divided into four strands. The meaning-focused input strand involves opportunities for learning through listening and reading where learners are already familiar with 95–98% of the running words in the listening and reading input. The meaning-focused output strand involves opportunities for learning through speaking and writing. The language-focused learning strand involves giving deliberate attention to vocabulary learning through activities such as intensive reading, using word cards, preteaching of vocabulary before doing communicative activities, and the deliberate learning of vocabulary strategies. The fourth strand is the fluency development strand where learners do not learn new vocabulary but practice making the best use of words

they already know. Each of these four strands deserves roughly equal time, and most of the remainder of this description of the role of vocabulary in language education will focus on how each of these strands can be put into practice.

Learning Vocabulary Through Listening

The conditions needed for the meaning-focused input strand are that learners should be focused on understanding what they are listening to or reading, they should already know at least 95% of the words and phrases in the input, and thus there is a small number of unfamiliar language features that can be understood through context clues. This means that at each level of a learner's language-proficiency development there needs to be material at an appropriate language level. Thus, it is essential to use specially written or simplified material where the vocabulary level is controlled in a planned way. An enormous amount of such material already exists in English, largely in the form of graded readers. Some of these can be used for listening activities in much the same way as a story is presented in weekly or daily installments on the radio. Another important source of listening input is the teacher's classroom instructions and interaction with learners. If these are carried out using high frequency vocabulary, then they become an important opportunity for vocabulary learning.

Most graded readers are fiction, and in addition to these, it is important to have nonfiction text as input. A useful activity that suits nonfiction material and that can help vocabulary learning is information transfer (Palmer, 1982). In such activities, listening input is changed into a diagrammatic or semi-diagrammatic form such as a labeled flow diagram or a completed information table. For example, the teacher describes the journey taken by a group of people and learners mark the route on a map. Similarly, while listening to a physical description of say a flowering plant, the learners complete a four-part table by filling in the names of the parts of the plant, where the parts are located, what they are like, and what job they do.

Another important source of vocabulary learning through meaning-focused input is through conversational interaction, and this will be looked at when we look at vocabulary learning through speaking.

Learning Vocabulary Through Reading

A well-designed, well-managed extensive reading program is an effective way of building vocabulary knowledge. It is important to realize that there are various kinds of vocabulary learning going on in extensive reading, representing different strengths of knowledge. An experiment by Waring and Takaki (2003) used three different kinds of vocabulary tests for each of the target words in a graded reader – (1) a recognition test (Which of these words occurred in the text?), (2) a multiple-choice test of word meaning, and (3) a translation test. The most difficult test, the translation test, showed small amounts of learning, but the other two tests showed increasingly larger amounts of learning. Thus, extensive reading provides opportunities for various degrees of learning for different words.

A good extensive reading program has the following features (Nation and Wang, 1999). It provides plenty of interesting material where each learner can read already knowing about 98% of the running words in the text. It encourages learners to read at least one book every two weeks. It encourages learners to read at least three books at the same level before moving up to the next level. It encourages learners to deliberately learn most of the general purpose unknown vocabulary that is met in the texts. It provides opportunities for learners to talk and write about their reading, being careful not to let reporting activities have the effect of reducing the amount of time spent reading. Extensive reading programs have many beneficial effects in addition to increasing vocabulary knowledge (Elley, 1991).

Deliberate Vocabulary Learning

All research comparing deliberate, decontextualized vocabulary learning with vocabulary learning from context has found that given the same amount of time, deliberate learning always results in more learning. This, however, should not be seen as an argument against learning from context but as part of the argument in favor of having deliberate learning as a component of a well-balanced course. In this part of a course, learners should study vocabulary using word cards (small cards with the target word or phrase on one side and its first language translation on the back). Although initially this kind of learning can be done in class time, it is best if it is done largely out of class when the learners have a few minutes of free time such as when traveling on the subway, waiting for a bus, or during TV commercials. Research provides useful guidelines for such learning.

1. Use small cards, not notebooks. Using cards allows the word or phrase to be put on one side and the translation on the other, which allows retrieval to occur, that is, looking at the word and then attempting to retrieve its translation from memory. Seeing both words together does

not allow the possibility of retrieval. Using cards also allows the order of the words to be easily changed to avoid serial learning.

2. Increasingly space the repetitions. That is, go through the pack of cards once and then do something else for five or ten minutes then go through the pack again. Then wait for half an hour before going through them again. Then wait two or three hours, and so on.
3. Use memory techniques such as the keyword technique (Nation, 2001: 311–314), word part analysis, study of etymologies, and visualization with the words that are difficult to remember. Skillful use of a dictionary to find a word's frequency of occurrence, to work out the core meaning of the word, and to see its range of senses, typical uses, and various forms also helps learning.
4. Make sure that the words in a pack do not have closely related meanings, that is, that there are not pairs of opposites or synonyms, or members of a lexical set such as articles of clothing, fruit, days of the week, or numbers. Learning such related items together makes learning 50% to 100% more difficult (Nation, 2000).
5. Say the words or phrases to yourself when looking at the cards and get help with the pronunciation of difficult words. Pronounceability is a major factor affecting vocabulary learning.
6. It is good to use translations of words when learning their meanings. The first language translation is a simple, clear way of communicating the meaning of a word, and research has shown such translations to be effective. As learners' proficiency develops, they will gain more accurate and elaborated representations of meaning.

Deliberate learning should become the responsibility of each learner, but learners may need encouragement and training in applying the guidelines.

Deliberate Vocabulary Teaching

The most obvious place for direct vocabulary teaching is as part of intensive reading. Intensive reading involves the teacher and the learners working carefully through a short text looking at a range of language features and content issues relevant to the text. From a vocabulary perspective, dealing with unknown words in intensive reading can involve preteaching some words before reading the text, quickly giving the meaning of some words, ignoring some, spending a lot of time on others, using the text as an opportunity to give training in the strategies of guessing from context, word-part analysis, and dictionary use, and altering the text to get rid of some low frequency words. The decision about which of these things to do for a particular word should depend on whether it is a high frequency word or a low frequency word, its importance for the message of the text, and the nature of the word itself and its context in the text. For example, preteaching should be done with high frequency words that are important for the message of the text, and that are not easily guessed from context. Low frequency words do not deserve this kind of attention in teaching. Altering the text should be a way of getting rid of low frequency words that are not important for the message of the text.

A small amount of deliberate teaching can occur in listening activities, and this has been shown to have a positive effect on vocabulary learning (Elley, 1989). Feedback on learners' spoken and written output is yet another opportunity for deliberate vocabulary teaching.

Deliberate vocabulary teaching also includes training learners in the use of the vocabulary strategies of guessing from context, learning from word cards, word part analysis, and dictionary use. These strategies require a repeated investment of time over several months so that learners reach the stage where it is easier to use the strategy than not to use it. For the guessing strategy, this can involve focusing one by one on the types of clues that are available from the immediate context and the wider context, having learners work with the teacher, then in pairs, and finally alone, and working with time pressure to develop fluency in using the guessing strategy. These strategies deserve a lot of classroom time because they can be used to cope with thousands of low frequency words

Learning Vocabulary Through Speaking

The meaning-focused output strand of a course involves learning vocabulary through speaking and writing. One of the most powerful factors affecting vocabulary learning through speaking is negotiation of meaning. Negotiation occurs when learners deliberately explain language items as a result of a breakdown in communication. The likelihood of negotiated words being learned is high, but negotiation still only accounts for a small proportion of the words learned through spoken communication activities. Most vocabulary learning in such activities occurs through guessing from context.

The way speaking tasks are designed can have a major influence on what and how much vocabulary is learned through such tasks. Vocabulary learning is helped if the written input to the task contains the target vocabulary, if performing the task requires the

use of that vocabulary, if there are repeated opportunities in the task for the vocabulary to be used, and if the task requires the learners to use the vocabulary in at least a slightly different way from the one in which it occurred in the written input to the task. It is not difficult to design speaking tasks that meet these requirements.

Learning Vocabulary Through Writing

Communication through the Internet has increased the opportunities for learning vocabulary through writing. There is evidence that written Internet discussion can include negotiation of vocabulary.

Productive information-transfer activities where learners turn diagrammatic representations such as graphs, plans and maps, and lists into connected written text provide good opportunities for vocabulary learning. The same design features described for speaking also apply in the design of writing tasks. Other writing activities that make use of written input such as synthesizing information from several texts can also be important sources of vocabulary learning.

In general, useful writing activities will involve the use of written or spoken input in getting information to write about.

Feedback on written work may also be an effective way of expanding vocabulary knowledge and collocational knowledge (which words go with which words). This feedback can be directed toward vocabulary choice and collocations.

Fluency Development

The fourth strand of a course is its fluency development strand. The vocabulary goal of this strand is to help learners make the best use of the vocabulary that they already know. Activities that fit within this strand do not aim to teach new words but aim to strengthen and enrich knowledge of words that are already partially known. Fluency activities have the following characteristics. They involve no unknown vocabulary, grammatical features, or discourse features. There is pressure or encouragement during the activity to perform at a faster than typical speed. The activities are message focused in that the learners aim to produce or comprehend messages, and, finally, they involve a large quantity of language use. Fluency activities are needed in each of the skills of listening, speaking, reading, and writing.

At the word or phrase level, fluency activities can focus on numbers, time sequences (yesterday, tomorrow, last week), greetings, and any set that needs to be used without a great deal of thought. In such activities, the teacher says words or phrases quickly while the learners point to what is being said. Such activities can also be done productively with the teacher pointing at pictures or symbols and the learners having to quickly produce the appropriate phrase. Other more advanced fluency activities include the 4/3/2 activity (giving the same talk in a decreasing timeframe to a new listener), speed reading, and ten-minute writing where learners write each day for ten minutes concentrating on quantity of writing rather than on quality.

Monitoring and Encouraging Progress

The learning of any individual word needs to be seen as a cumulative process with both strength of knowledge and richness of knowledge of each word growing with repeated receptive and productive encounters with a word. Because there are many aspects to learning a word and many degrees of strength of knowledge, there is a wide variety of vocabulary measures available, and each one needs to be seen as a means of providing at least a slightly different view of learners' vocabulary knowledge (Read, 2000). The choice of a vocabulary test item needs to be based on a careful consideration of what kind of knowledge needs to be measured. Factors to consider when choosing a particular test format include, (1) the kind of knowledge that needs to be tested, (2) the likely strength of the knowledge (using a demanding test to test knowledge that is likely to be partial or weak will not provide a useful measure), (3) the time available and the number of items that there should be in the test, and (4) the language proficiency and test-making skills of the test makers and test markers. To reach an adequate degree of reliability, most vocabulary tests need to have at least 30 items.

Well-researched vocabulary measures include a variety of vocabulary-richness measures which look at the range of vocabulary used in a piece of writing (Malvern *et al.,* 2004), interview measures that explore knowledge of the meanings of words and their use (Nagy *et al.,* 1985; Wesche and Paribakht, 1996), multiple-choice depth measures which look at collocations and meaning components (Read, 1998), and multiple-choice, matching, and translation items which focus on word meanings.

Vocabulary testing can have a variety of goals. Because vocabulary knowledge is such a fundamental part of language proficiency, vocabulary tests have often been used for wider goals.

1. Placement in a language program. Measures such as the Vocabulary Levels Test (Schmitt *et al.,*

2001), the Eurocentres 10KA (Meara and Jones, 1990), and the vocabulary levels dictation test (Fountain and Nation, 2000) have often been used as part of a wider battery of tests to assign learners to the appropriate proficiency level group in an intensive English program.
2. Diagnosis of vocabulary strengths and weaknesses. The Vocabulary Levels Test was designed to see what levels of vocabulary learners already knew and to indicate the next level to focus on in teaching and learning.
3. Encouragement of learning. Some programs use weekly vocabulary tests to encourage learners to do deliberate vocabulary learning. These tests can involve each learner giving the teacher a list of ten of the new words they have been working on, and then the teacher writes a symbol next to each word indicating what the learner has to do with each word in the test. For example, S can mean write a sentence containing that word, C can mean write two collocations for the word, M can mean explain the meaning of the word, and so on (Smith, 1996).
4. Monitoring progress. Vocabulary tests that are focused on the learning done in a course (achievement tests) may be used to measure progress through the course.
5. Measuring proficiency. Measures of total vocabulary size, and measures of particular vocabulary levels such as the Academic Word List, may be used to see where learners are in their vocabulary growth.

Encouraging Autonomy

Like most learning, vocabulary learning will be most effective if learners take control of and responsibility for their learning. This involves knowing what to learn, knowing how to learn it, and being motivated to do this learning and to put it to use.

This is not easy to achieve, and the few studies on vocabulary in this area indicate that most learners do not take an organized approach to their learning (Moir and Nation, 2002). One way of encouraging autonomy is to introduce a negotiated syllabus where the teacher and learners share the decision making about the various aspects of the course on a continuing basis. These aspects may include what vocabulary to learn, what vocabulary activities to do in class and how much time to spend on them, what vocabulary tests to use, and what homework to set. The goals of a negotiated syllabus are to make the course as sensitive as possible to the changing needs of the learners and to get the learners to feel ownership of the course.

Principles of Vocabulary Learning and Teaching

The vocabulary component of a language course will be more effective if it is based on well-supported principles that are clearly known by both teachers and learners. A short list of such principles should include the following.

1. The sequence of vocabulary learning should move from high frequency vocabulary and special purposes vocabulary to low frequency vocabulary.
2. High frequency vocabulary and special purposes vocabulary should get attention across the four, roughly equal, strands of teaching and learning vocabulary: meaning-focused input, meaning-focused output, language-focused learning, and fluency development.
3. With low frequency vocabulary, teachers should focus on the strategies of guessing from context, learning word cards, using word parts, and using the dictionary.
4. Learning activities should be designed to encourage thoughtful processing of vocabulary through retrieval, generative use, and the use of mnemonic devices where needed.
5. Learners should be helped to take responsibility for their own vocabulary learning.

Vocabulary growth is a critical aspect of second language proficiency development, and as such it needs to be properly supported by a principled, well-thought-out component in a language course.

See also: Corpus Studies: Second Language; Lexical Acquisition; Second and Foreign Language Learning and Teaching; Second Language Listening; Second Language Reading; Second Language Speaking; Second Language Writing; Vocabulary: Second Language.

Bibliography

Coxhead A (2000). 'A new academic word list.' *TESOL Quarterly 34,* 213–238.

Elley W B (1989). 'Vocabulary acquisition from listening to stories.' *Reading Research Quarterly 24,* 174–187.

Elley W B (1991). 'Acquiring literacy in a second language: the effect of book-based programs.' *Language Learning 41,* 375–411.

Fountain R L & Nation I S P (2000). 'A vocabulary-based graded dictation test.' *RELC Journal 31,* 29–44.

Malvern D, Richards B, Chipere N & Duran P (2004). *Lexical diversity and language development.* Basingstoke: Palgrave Macmillan.

Meara P & Jones G (1990). *Eurocentres vocabulary size test 10KA.* Zurich: Eurocentres.

Moir J & Nation I S P (2002). 'Learners' use of strategies for effective vocabulary learning.' *Prospect 17,* 15–35.
Nagy W E, Herman P & Anderson R C (1985). 'Learning words from context.' *Reading Research Quarterly 20,* 233–253.
Nation I S P (2000). Learning vocabulary in lexical sets: dangers and guidelines. *TESOL Journal 9,* 6–10.
Nation I S P (2001). *Learning vocabulary in another language.* Cambridge: Cambridge University Press.
Nation P & Wang K (1999). 'Graded readers and vocabulary.' *Reading in a Foreign Language 12,* 355–380.
Palmer D M (1982). 'Information transfer for listening and reading.' *English Teaching Forum 20,* 29–33.
Read J (1998). 'Validating a test to measure depth of vocabulary knowledge.' In Kunnan A J (ed.) *Validation in language assessment.* Mahwah, NJ: Lawrence Erlbaum. 41–60.
Read J (2000). *Assessing vocabulary.* Cambridge: Cambridge University Press.
Schmitt N & McCarthy M (eds.) (1997). *Vocabulary: description, acquisition and pedagogy.* Cambridge: Cambridge University Press.
Schmitt N, Schmitt D & Clapham C (2001). 'Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test.' *Language Testing 18,* 55–88.
Smith H (1996). 'An individualised vocabulary programme.' *TESOLANZ Journal 4,* 41–51.
Waring R & Takaki M (2003). 'At what rate do learners learn and retain new vocabulary from reading a graded reader?' *Reading in a Foreign Language 15,* 130–163.
Wesche M & Paribakht T S (1996). 'Assessing second language vocabulary knowledge: depth versus breadth.' *Canadian Modern Language Review 53,* 13–40.

Language Families and Linguistic Diversity

M Ross, The Australian National University, Canberra, Australia

Linguistic Diversity

'Linguistic diversity' refers to several interrelated phenomena:

1. phylogenetic (genetic, genealogical) diversity: the number of language families in a geographic area;
2. intrafamilial and intralinguistic variation: differences among the languages of a family and among the variants of a language;
3. language density: the number of languages in a geographic area;
4. typological diversity: differences among the structures which make up languages.

The focus of this article is on (1), (2) and (3) (*see also* **Variation and Language: Overview**). More conventional approaches to these topics are discussed first, then a recent alternative.

Although (4) will not be discussed further here, the relationship between phylogenetic and typological diversity should be mentioned. It is quite common for the members of a language family to be typologically similar, but there is no necessary reason why they should be so. There is, for example, considerable typological variety among the grammars of the Austronesian family, some of it the result of contact with non-Austronesian languages, but much of it the outcome of language-internal changes. Conversely, languages belonging to different language families may be typologically quite similar, either as the result of contact or as the result of independent parallel innovation.

Language Families

Definitions and Problems

A language family is conventionally defined as a set of languages that share a common ancestor. The family metaphor captures two insights: that languages are systems that are transmitted from one generation to the next (generational continuity) and that these systems change over time so that when a community of speakers divides into, say, three separate communities, the speech of each of the three will change in different ways from the others, leading eventually to mutual unintelligibility, i.e., three languages descended from a shared ancestor. The basis of the metaphor is biological evolution, where one species divides into two or more new species (McMahon, 1994: Chap. 12), even though linguists often use the term 'family tree,' which otherwise denotes trees drawn to represent human family relationships.

Identifying a language family in the first place depends on finding individual-identifying evidence (Nichols, 1996, 1997), patterned similarities between a set of languages that could not have arisen by chance and must be the outcome of shared inheritance. The Indo-European family, for example, was first definitively identified by Sir William Jones in 1786 on the basis of similarities between the verb paradigms of Sanskrit, Greek, and Latin. Recently, the Trans-New

neurolinguistics (1976–1979). Another series of edited books on neurolinguistics started in 1973. As in the 19th century, rapid developments in the area of aphasia triggered similar developments in the study of other cognitive disorders following brain disorders.

Bibliography

Arbib M, Caplan D & Marshall J C (1982). 'Neurolinguistics in Historical Perspective.' In Arbib M, Caplan D & Marshall J C (eds.) *Neural models of language process.* San Diego: Academic Press.
Bleser R de (1987). 'From agrammatism to paragrammatism: German aphasiological traditions and grammatical disturbances.' *Cognitive Neuropsychology 4,* 187–256.
Caplan D (1994). 'Language and the brain.' In Gernsbacher M A (ed.) *Handbook of psycholinguistics.* San Diego: Academic Press.
Caplan D (1987). *Neurolinguistics and linguistic aphasiology: an introduction.* New York: Cambridge University Press.
Eling P (ed.) (1994). *Reader in the history of aphasia.* Amsterdam: Benjamins.
Finger S (1994). *Origins of neuroscience: a history of explorations into brain function.* Oxford: Oxford University Press.
Glozman J (1996). 'A brief history of Russian aphasiology.' *Neuropsychology Review 6,* 1–10.
Goodglass H & Wingfield A (1988). 'The changing relationship between anatomic and cognitive explanation in the neuropsychology of language.' *Journal of Psycholinguistic Research 27,* 147–165.
Goodglass H (1988). Historical perspective on concepts of aphasia. In Boller F & Grafman J (eds.) *Handbook of neuropsychology,* vol. 1. Amsterdam: Elsevier.
Lecours A R, Chain F, Poucet M, Nespoulous J L & Joanett Y (1992). 'Paris 1908: The hot summer of aphasiology or a season in the life of a chair.' *Brain and Language 42,* 105–152.
Lichtheim H (1885). 'On aphasia.' *Brain 7,* 433–485.
Noppeney U & Wallesch (2000). 'Language and cognition. Kurt Goldstein's theory of semantics.' *Brain and Cognition 44,* 367–386.
Penfield W & Roberts L (1959). *Speech and brain mechanisms.* New Jersey: Princeton University.
Tesak J (2001). *Geschichte der aphasie.* Idstein: Schulz-Kirchner Verlag.
Weisenburg T & McBride K (1935). *Aphasia: a clinical and psychological study.* New York: The Commonwealth Fund.
Wernicke C (1874). *Der Aphasichse Symptomencomplex. Eine psychologische Studie auf anatomischer Basis.* Breslau: Cohn & Weigert.
Whitaker H & Whitaker H A (1976–1979). *Studies in Neurolinguistics* (Vols 1–4). New York: Academic Press.
Whitaker H A (1988). 'Neurolinguistics from the middle ages to the premodern era: historical vignettes.' In Stemmer B & Whitaker H A (eds.) *Handbook of neurolinguistics.* San Diego: Academic Press. 27–54.
Zurif E & Swinney D (1994). 'The neuropsychology of language.' In Gernsbacher M A (ed.) *Handbook of psycholinguistics.* San Diego: Academic Press.

Language Education and Gender

B Egbo, University of Windsor, Windsor, Ontario, Canada

Research in language and gender as a subfield in sociolinguistics has accelerated during the last several decades largely due to the critical work and writings of feminist sociologists, anthropologists, sociolinguists, and post-structuralists (e.g., Lakoff, 1975; Philips *et al*., 1987; Coates, 1993), much in the same way as Simone de Beauvoir's seminal work *The second sex*, released in 1953, galvanized contemporary feminist movement. Emerging naturally from the interest in language and gender have been inquiries that collectively cohere around a critical analysis of the interface between language, education, and gender, which is the subject of this review.

Different from its classificatory or grammatical meaning, gender is examined here in terms of a social construct that society arbitrarily ascribes to men and women, based on their perceived roles in society. Such a framework recognizes that while the biological state of 'femaleness' and 'maleness' is, recent claims to the contrary notwithstanding (see Bergvall *et al*., 1996), stable and universal, understandings of gender are in a constant state of flux, temporal, depend on the prevailing social order, and mediate the construction of the gendered self, which in turn, shapes individual identity (Bjerrum Nielsen and Davies, 1997; Egbo, 2004).

While it is not clear how it happens, given the polemic among sex-role development theorists, researchers agree that cross-culturally, by the time children begin their formal schooling, they have developed an emotional attachment to their gender, and already act, speak, and behave according to conventional images of gender – behaviors and discourse norms that are further reified in schools (Spender, 1982; Bjerrum Nielsen and Davies, 1997).

Across all educational spectrums, language as a complex technology of communication serves two diametrically opposing functions in addition to, of course, its functional uses in teaching and learning. First, it enhances linguistic and communicative competence in a Habermian sense (see Habermas, 2001), through formal and informal discursive interchange among the various actors. Second, it serves a gatekeeping function *vis à vis* the safeguarding of the privileged and highly valued cultural and linguistic capital of powerful groups in society, such as men, which enables them to preserve their dominance. Pierre Bourdieu's (1991) sociological account of the ways through which educational systems reproduce the differential power relations that exist in wider society is at once a critique of educational practices and a recognition of language as a symbolic tool for domination. However, a third and emergent view, which perhaps evolved from the postulates of critical theory and is gaining recognition among feminists and critical pedagogists, sees a more dialectical and transformative role for language in education than its hegemonizing power. There is now a widely held view among proponents of critical language practices that if schools' discourse norms contribute to the maintenance of the gender order in society, the same discursive practices also hold emancipatory possibilities which through critical interrogation and agency should, as the argument goes, lead to consequential change in wider society (Corson, 1993; Blackmore and Kenway, 1995).

Recurrent Debates about Language, Education, and Gender

In this section, I examine three recurrent themes in the ongoing debate on language, education, and gender. It should be noted from the onset that while these three strands of the debate are considered separately here for conceptual clarity, the distinctions are less discrete in practice.

Three Strands of the Debate

Beginning in the mid-1960s, one major research focus in the area of language, education, and gender, has been challenging the orthodoxy of the taken-for-granted canonical knowledge that is disseminated in schools including language practices and discourse norms. The evidence assembled from early research in the field suggested that androcentric biases in wider society had been transplanted into the classroom (Spender, 1982; Kelly, 1988; Stanworth, 1987, for example) despite the simultaneous collapse of gender differentiation in education and the emergence of the equal rights movement in the 1950s (Bjerrum Nielsen and Davies, 1997). The school curriculum reflected male world views including male linguistic register, while women's discourse norms were devalued, despite the evidence (although inconclusive in some areas) that have pointed to gender-distinctive language use and linguistic practices (Lakoff, 1975; Philips *et al*., 1987; Tannen, 1991; McCormick, 1994). This was true of Western societies as it was true of non-Western colonial and post-colonial societies (Egbo, 2004). Also problematic was the issue of gender representations in educational texts (including narratives, drama, poetry, etc.).

The 1970s marked a turning point, however. Post-structural feminists' challenges to the prevailing orthodoxy, their engagement of dominant texts which aimed at unmasking the trajectories between knowledge and power on the one hand and discourses as sites of human struggle on the other, resulted in policies and practices that were aimed at leveling the 'discursive' field. The language of education and the classroom in particular took a seemingly inclusive turn – a metamorphosis that also shifted to some degree the language of texts and textbooks, which became less stereotyped, paternalistic, more nuanced and neutral. Far from settling the matter, however, some writers have argued that the linguistic neutrality that now pervades educational texts simply masks the continued dominance of male discourse norms under the "cloak of egalitarian educational discourse" (Bjerrum Nielsen and Davies, 1997: 126).

In contrast to the views outlined above, some researchers have expressed reservations about uncritical claims of gender bias in texts because cognitively and affectively speaking, individual engagement with text is a subjective experience (Sunderland, 2000; Swann, 2003), albeit not to the point of disinheriting one's cultural capital in the process (Balfour, 2003). With particular reference to language education, Sunderland argues compellingly that it is difficult to prove gender bias in language texts because one cannot accurately predict a reader's reaction to the text.

A second research focus has been the nature of classroom discourse and teacher-student talk. The bulk of the research on this issue converges around the following issues: pedagogical practices and classroom discourse norms tend to silence girls, teachers call on boys more often than girls, boys talk more often and initiate dialogue with teachers who act as enablers: the overwhelming majority of classroom talk is either directed at or dominated by boys (Spender, 1982; Stanworth, 1987; Graddol and Swann, 1989). With regard to classroom talk, the nexus of the various arguments is that boys'

dominance and assertiveness are extensions of the unequal power relations and exclusionary language practices in wider society in which women are positioned as the "other," while men are accorded "autonomous and varied linguistic status" (Thorne *et al.*, 1983: 9).

A third issue of interest has been the nature of gender-based teacher-student interactions. A number of classroom observations have catalogued ways in which teachers interact differentially with each gender. Boys appear to have more interactions with teachers, are called upon more often, and generally have more contact time with teachers. Teachers are also said to be more favorably disposed to boys' participation in class (Shakeshaft, 1986; Sadker and Sadker, 1986; Stanworth, 1987; Kelly, 1988).

There are, however, contradictory views that suggest that the disproportionate attention that boys receive relative to girls may be a collateral consequence of behavior-related reprimands, rather than deliberate discriminatory action on the part of teachers (Sunderland, 2000). Moreover, studies of classroom social interactions have shown that both genders do in fact sometimes engage in heterodoxical behavior (Gallas, 1998; Sheldon, 1997). For example, on the strength of the evidence from an in-depth study of her own classroom, Gallas (1998) believes that girls sometimes traverse gender borders, put up resistance, and generally act in unpredictable ways. She further argues that while stereotypes may exist, heterogeneity may be more reflective of classroom interactions than is normally assumed.

Cumulatively, and as a practical matter, however, gender-biased language practices in education can have psychosocial and material consequences for girls. While there have been significant reductions in the use of sexist language in educational settings, empirical research relating to gender and education continues to cite sex-role stereotyping and unfair language practices as some of the factors that mitigate against girls and women at all levels of schooling. Based on a review of previously assembled evidence, Corson (1993) argues that social identities are constructed through school texts and other educational materials, all of which position and define individuals in relation to knowledge and power. With particular reference to gendered language practices, he argues that: "practices and policies found in the objective discursive structures of schooling itself, help to create and reinforce disadvantages for girls and women" (Corson, 1993: 139). In the same vein, Luke (1994: 363) calls for the de-gendering of literacy activities because the sexual division of literacy "plays into the patriarchal reproduction of knowledge, competence, and, ultimately, economic structure."

In short, gendered language practices can impact the overall life chances of girls and women. In many parts of the world, access to certain kinds of language and discourse is often associated with increases in life chances which are a function of two elements – options and ligatures. *Options* are choices, while *ligatures* are bonds that individuals form through immersion in a given culture or by virtue of their social positions and roles in society (Dahrendorf, 1979). Options and ligatures are critical to the empowerment of any group, particularly those who have been historically oppressed. Cross-culturally, gendered language practices in schools are implicated in the limited options girls and women have in their respective societies.

Socio-Politico-Cultural Variations

Studies of gendered language practices and linguistic covariation in contexts other than Western societies, from where the bulk of current empirical data originates, are few and far between. However, as pointed out above, culture has significant influences on gender differentiated language uses and practices in many societies (see Philips *et al.*, 1987). Emergent knowledge from non-Western societies reaffirm prior evidence from Western contexts that gendered language practices in wider society have corresponding effects on language and discourse practices in schools (Balfour, 2003). For example, in some societies, women are not expected nor indeed allowed to speak in the public sphere, on ideological or religio-social grounds. This means that in public and quasi-public domains such as schools, girls must maintain a culture of silence. In other societies, socially sanctioned gender-typed differential treatment from teachers is the norm rather than the exception (see Kitetu and Sunderland, 2002). While such treatment may appear discriminatory at first glance, in some cultural contexts, according to Kitetu and Sunderland, differentiated gendered discourse norms in schools may be based on societal valuation of difference and the desire to protect women's interests, thus adding another layer of confusion to an already complex matter.

Challenges Ahead

Given the persistence of gender-biased language and linguistic practices in educational settings, a key question becomes what should be the future directions of research and praxis? Pathways to more empowering experiences for girls and women *vis à vis* language practices in education depend first and foremost on a critical linkage between theory and practice.

A second line of intervention should begin at teacher preparation institutions, where as part of their

curriculum novice teachers are taught the art of doing critical language and pedagogical analysis of texts based on the concept of critical discourse analysis (Fairclough, 1995). Under this arrangement, novice teachers learn ways of deconstructing and engaging the nonneutral properties of texts and discourses, seeing language instead as a dialogical interchange between social actors (Bakhtin, 1981).

Related to, but different from, the idea of forging a link between theory and practice is the issue of new research directions in the field. While a shift is already in progress, critical praxis will also depend on research that advances our understandings of the trajectories of language, education, and gender beyond the dichotomizing binary thinking that result in essentialist oversimplifications of a relationship that is far from straightforward (Swann, 2003; Sunderland, 2000; Tannen, 1996; Bergvall *et al.*, 1996). There is indeed an amalgam of interacting variables (e.g., ethnicity, social class, religion, and personality) that impinge on the construction of the gendered self, which in turn influences linguistic behavior and practices in and out of the school. In effect, rigid adherence to deterministic assumptions of masculinity and femininity *vis à vis* language and discourse norms in education, may be as conceptually limiting as the unwavering adherence to the dominant male linguistic orthodoxy that sustains the gender divide in society in the first place.

See also: Classroom Talk; Discrimination and Language; Gender and Language; Gender, Sexuality and Language; Sexist Language; Variation in Pidgins and Creoles.

Bibliography

Bakhtin M M (1981). *The dialogic imagination.* Austin, TX: University of Texas Press.

Balfour R (2003). 'Between the lines: Gender in the reception of texts by school children in rural KwaZulu-Natal, South Africa.' *Gender and Education 15(2),* 183–198.

Bergvall V L, Bing J M & Freed A F (eds.) (1996). *Rethinking language and gender research: Theory and practice.* London: Addison Wesley Longman.

Bjerrum Nielsen H & Davies B (1997). 'The construction of gendered identity through classroom talk.' In Davies B & Corson D (eds.) *Encyclopedia of language and education vol. 3: Oral discourse and education.* Dordrecht: Kluwer Academic Publishers. 125–135.

Blackmore J & Kenway J (1995). 'Changing schools, teachers, and curriculum: What about the girls?' In Corson D (ed.) *Discourse and power in educational organizations.* Toronto: OISE Press. 233–256.

Bourdieu P (1991). *Language and symbolic power.* Cambridge: Polity Press.

Coates J (1993). *Women, men and language: A sociolinguistic account of gender differences in language* (2nd edn.). London: Longman.

Corson D (1993). *Language, minority education and gender: linking social justice and power.* Clevedon: Multilingual Matters.

Dahrendorf R (1979). *Life chances.* Chicago: University of Chicago Press.

Egbo B (2004). 'Intersections of literacy and construction of social identities.' In Pavlenko A & Blackledge A (eds.) *Negotiation of identities in multilingual contexts.* Clevedon: Multilingual Matters. 243–265.

Fairclough N (1995). *Critical discourse analysis.* London: Longmans.

Gallas K (1998). *Sometimes I can be anything: Power, gender and identity in a primary classroom.* New York: Teachers College Press.

Graddol D & Swann J (1989). *Gender voices.* Oxford: Basil Blackwell.

Habermas J (2001). *On the pragmatics of social interaction.* Cambridge, Mass: MIT Press.

Kelly A (1988). 'Gender differences in teacher–pupil interactions: A meta-analytic review.' *Research in Education 39,* 1–23.

Kitetu C & Sunderland J (2002). *Gendered discourses in the classroom: The importance of cultural diversity.* CRILE Working Paper, Lancaster University.

Lakoff R (1975). *Language and women's place.* New York: Harper and Row.

Luke A (1994). 'On reading and the sexual division of literacy.' *Journal of Curriculum Studies 26,* 361–381.

McCormick K M (1994). 'Gender and language.' In Asher R E & Simpson J M (eds.) *The Encyclopedia of Language and Linguistics,* vol. 3. Oxford: Permagon Press. 1353–1360.

Philips S, Steele S & Tanz C (eds.) (1987). *Language, gender and sex in comparative perspectives.* Cambridge: Cambridge University Press.

Sadker M & Sadker D (1986). 'Sexism in the classroom: From grade school to graduate school.' *Phi Delta Kappan 67,* 512–515.

Shakeshaft C (1986). 'A gender at risk.' *Phi Delta Kappan 67,* 499–503.

Sheldon A (1997). 'Talking power: Girls, gender enculturation and discourse.' In Wodak R (ed.) *Gender and Discourse.* London: Sage. 225–244.

Spender D (1982). *Invisible women: the schooling scandal.* London: Writers and Readers Publishing Cooperative Society.

Stanworth M (1987). 'Girls on the margins: A study of gender divisions in the classrooms.' In Weiner G & Arnot M (eds.) *Gender under scrutiny: New inquiries in education.* London: Unwin Hyman. 198–212.

Sunderland J (2000). 'New understandings of gender and language classroom research: Texts, teacher talk and student talk.' *Language Teaching Research 4,* 149–173.

Swann J (2003). 'Schooled language: Language and gender in educational settings.' In Holmes J & Meyerhoff M (eds.) *The handbook of language and gender.* Malden, MA: Blackwell Publishing. 624–644.

Tannen D (1991). 'Teachers' classroom strategies should recognize that men and women use language differently.' *The Chronicle of Higher Education 37(40),* B2–3.
Tannen D (1996). 'Researching gender related patterns in classroom discourse.' *TESOL Quarterly 30,* 341–344.
Thorne B, Kramarae C & Henley N (1983). 'Language, gender and society: Opening a second debate of research.' In Thorne B, Kramarae C & Henley N (eds.) *Language, gender and society.* Rowley: Newbury House Publishers. 7–24.

Language Education for Endangered Languages

L Grenoble, Dartmouth College, Hanover, NH, USA

As local communities become more aware of the potentially imminent loss of their languages, they have begun to undertake measures to foster education in them. There is currently a wide range of programs which have been adopted to arrest language loss and to help children and adults take control of learning their language. These can loosely be divided into school-based and community-based programs.

School-Based Programs

When people think of education, they naturally think of school programs, and perhaps the majority of programs for endangered languages are based in schools. In many communities the school is seen as the primary vehicle for language education. Note that this is a radical shift from the way that children first learn their primary language, within the home and within a community of speakers. It is more reminiscent of the way that children typically learn foreign languages.

Total-Immersion Programs

Total-immersion programs work well when there is a sufficient speaker base to help insure their success. These programs tend to be multifaceted, with total-immersion language instruction in the schools reinforced by the sole use of the language in the home. Thus they are often most successful when fluent speakers can be found in all generations, as the source not only of teachers in the school, but also of parents who can speak the language in the home. A common goal of total-immersion programs is complete and fluent use of the target language in all domains. In a true total-immersion program, all of the school curriculum is taught in the endangered language and the language of wider communication is often taught as a secondary subject. Bi- or multilingualism is an ultimate goal, as in general, children enrolled in such programs ultimately need full use of both the endangered language and the language of wider communication.

The Language Nest Model

The language nest model is a particular type of immersion program. It was first used in the 1980s in New Zealand for the revitalization of the Maori language and is often referred to by its Maori name, Te Kōhanga Reo. It has subsequently also become closely associated with Hawaiian language revitalization, which built its program on the Te Kōhanga Reo model but uses the Hawaiian name, Pūnana Leo. The language nests were initially created in regions where the only group of people speaking the language was, by and large, the grandparent generation; the parents and children tend to be monolingual in the language of wider communication. The language nest model takes preschool children and places them in 'nests' where the endangered language is spoken. Parents are generally expected to make the commitment to learning the language and using it in the home to reinforce what is learned at school. Historically, as each lead class graduated and went on to a higher level, first in preschool, then elementary and secondary school, new nesting classes were created so that children ultimately received all of their education in the target language.

The creation of domains of language use outside of the schools has proven to be a challenge for language nest programs, in large part because the speaker base is primarily second-language speakers who do not attain full fluency in the endangered language.

Partial-Immersion Programs

For languages with relatively few remaining speakers or with a lower level of community commitment, it is often not realistic to implement a full-fledged immersion program. Partial-immersion programs exist in many schools, where the curriculum is taught primarily in the language of wider communication and the endangered language is taught in specially focused classes. There are two basic types of partial-immersion programs. In the one, the less common, some subjects are taught in the endangered language. Such subjects are typically the culture and/or history of the people. The benefits here are that learners acquire the language in culturally appropriate settings

Tannen D (1991). 'Teachers' classroom strategies should recognize that men and women use language differently.' *The Chronicle of Higher Education 37(40),* B2–3.
Tannen D (1996). 'Researching gender related patterns in classroom discourse.' *TESOL Quarterly 30,* 341–344.
Thorne B, Kramarae C & Henley N (1983). 'Language, gender and society: Opening a second debate of research.' In Thorne B, Kramarae C & Henley N (eds.) *Language, gender and society.* Rowley: Newbury House Publishers. 7–24.

Language Education for Endangered Languages

L Grenoble, Dartmouth College, Hanover, NH, USA

As local communities become more aware of the potentially imminent loss of their languages, they have begun to undertake measures to foster education in them. There is currently a wide range of programs which have been adopted to arrest language loss and to help children and adults take control of learning their language. These can loosely be divided into school-based and community-based programs.

School-Based Programs

When people think of education, they naturally think of school programs, and perhaps the majority of programs for endangered languages are based in schools. In many communities the school is seen as the primary vehicle for language education. Note that this is a radical shift from the way that children first learn their primary language, within the home and within a community of speakers. It is more reminiscent of the way that children typically learn foreign languages.

Total-Immersion Programs

Total-immersion programs work well when there is a sufficient speaker base to help insure their success. These programs tend to be multifaceted, with total-immersion language instruction in the schools reinforced by the sole use of the language in the home. Thus they are often most successful when fluent speakers can be found in all generations, as the source not only of teachers in the school, but also of parents who can speak the language in the home. A common goal of total-immersion programs is complete and fluent use of the target language in all domains. In a true total-immersion program, all of the school curriculum is taught in the endangered language and the language of wider communication is often taught as a secondary subject. Bi- or multilingualism is an ultimate goal, as in general, children enrolled in such programs ultimately need full use of both the endangered language and the language of wider communication.

The Language Nest Model

The language nest model is a particular type of immersion program. It was first used in the 1980s in New Zealand for the revitalization of the Maori language and is often referred to by its Maori name, Te Kōhanga Reo. It has subsequently also become closely associated with Hawaiian language revitalization, which built its program on the Te Kōhanga Reo model but uses the Hawaiian name, Pūnana Leo. The language nests were initially created in regions where the only group of people speaking the language was, by and large, the grandparent generation; the parents and children tend to be monolingual in the language of wider communication. The language nest model takes preschool children and places them in 'nests' where the endangered language is spoken. Parents are generally expected to make the commitment to learning the language and using it in the home to reinforce what is learned at school. Historically, as each lead class graduated and went on to a higher level, first in preschool, then elementary and secondary school, new nesting classes were created so that children ultimately received all of their education in the target language.

The creation of domains of language use outside of the schools has proven to be a challenge for language nest programs, in large part because the speaker base is primarily second-language speakers who do not attain full fluency in the endangered language.

Partial-Immersion Programs

For languages with relatively few remaining speakers or with a lower level of community commitment, it is often not realistic to implement a full-fledged immersion program. Partial-immersion programs exist in many schools, where the curriculum is taught primarily in the language of wider communication and the endangered language is taught in specially focused classes. There are two basic types of partial-immersion programs. In the one, the less common, some subjects are taught in the endangered language. Such subjects are typically the culture and/or history of the people. The benefits here are that learners acquire the language in culturally appropriate settings

using the specific lexicon. Thus a Native American class on oral traditions or crafts could be conducted in the indigenous language. In the other, the endangered language is taught as a secondary subject. The advantages to this model are obvious: it requires a smaller commitment of time and resources. Moreover, children coming to these programs often do not know the endangered language, and so need to learn it from the beginning. The primary disadvantage is that the endangered language is allotted secondary status within the schools and is taught more as a foreign language than as a primary language. Thus it does not achieve the full range of uses of the language of wider communication and cannot supplant it. Although such programs may be designed with the ultimate goal of stable bilingualism, they rarely achieve it, since they do not achieve full fluency in the endangered language.

Outside of the Schools

The dichotomy between school-based programs and others is somewhat artificial, as often two different types of programs are used in conjunction with each other. For example, many of the total-immersion school programs require the parents to learn the language and use it in the home to reinforce the work that is done in the schools. While school-based programs are almost exclusively built on Western models which require literacy in the endangered language, programs outside of the schools often dispense with literacy as too time-consuming and cumbersome for learning to communicate orally.

Community-Based Programs

Community-based programs are grassroots movements, stemming from and sustained by community leadership and commitment to language education. Such programs vary greatly in nature, but often have in common that they bring together a variety of community members for language practice. Maori communities in New Zealand organize weeklong or weekendlong immersion events for community members of all ages; people come together with the commitment to speak nothing but Maori for the time of the event. Other communities have established similar immersion events, or schedule regular meetings for language practice. These group meetings often link the practice of native crafts and traditions with language use. They have the advantage of offering options to adult learners who may not be able to attend regular classes. Beyond creating opportunities for language use and immersion, such programs organize language festivals, feature plays, songs, and poems, and actively promote the use of language in conjunction with traditional activities.

Some programs have made the decision to start teaching the endangered language to adults, not children. The rationale behind the decision is that it is the adults who will then use the language and teach it to the children, as is the norm for first-language acquisition. UNESCO, for example, advocates adult literacy first, with an emphasis on the acquisition of the practical skills needed for functioning in the modern workplace.

Master–Apprentice Program One very successful program for language groups with few remaining speakers is the master–apprentice program. Initiated in California in 1992 for teaching languages on the verge of extinction, its basic principles can be adopted to the teaching of endangered languages anywhere. The program matches an 'apprentice' learner with a 'master' of the language whose job is to teach the language by using it in everyday life. The program is based on the following principles: (1) the use of English is not permitted in interactions between master and apprentice; (2) the apprentice needs to be a full participant in determining the content of the program and assuring use of the target language; (3) oral, not written, language use is always primary in learning and communicating; (4) learning occurs not in the classroom, but in real-life situations, engaging in real-life activities (e.g., cooking, gardening, etc.); and (5) comprehension will come to the beginning language learner through the activity, in conjunction with nonverbal communication. These principles are designed to insure that language learning and instruction take place in an immersion setting that nearly replicates the 'natural' language-learning environment of children (as opposed to artificial classroom settings, for example).

Beyond the inherent incentive to learn and to teach one's language, in the Californian program a modest stipend is provided for up to 3 years of work between the master and the apprentice. After this period the graduate apprentices should be prepared to continue their education with the masters, but also to serve as teachers themselves.

Training before beginning the team work is critical. The language masters are primarily tribal elders who may not have actively used their language for many years, due to the very factors which have led to language attrition. In addition, they are not trained language teachers and many have never taught their language. Thus initial training sessions are designed to provide the opportunity for masters to become accustomed to speaking their languages again and to introduce the basic principles of language immersion

instruction, such as building and practicing vocabulary, and enforcing the importance of repetition, review, and patience in language learning. Important components of the training include getting the participants used to nonverbal communication, teaching apprentices key phrases and questions in the target language, and some cross-cultural comparison of the different ways language is used in different cultures. Despite the many linguistic and cultural differences, there is a commonality of experience which makes it very useful for all teams, from beginning to advanced, to come together for this training.

By the end of the first year, apprentices should be able to ask and answer simple questions about themselves, describe pictures, use some culture-specific language (prayers, stories, etc.), and recite a short speech prepared with the help of the master. This basic repertoire is expanded in the second year, with the goals of being able to speak in simple grammatical sentences, being able to carry on extended conversations, having increased comprehension, being able to converse on most topics, and being able to give short speeches. Finally, by the end of the third year of the program, the apprentices should be able to converse at length, use long (and presumably complicated) sentences, and develop plans for teaching the language. These goals are at once realistic and ambitious: language learning is a slow process, and the apprentice meets with the teacher only for 10 hours per week. One can predict that the learner's motivation will be very high in this program and will have a positive impact on learning. The results of the program vary among individual students depending on a range of factors, such as the overall time commitment, how much the apprentice is truly immersed in the language, and so on.

Potential Difficulties

Despite the many differences between these various educational models and the particularities of individual endangerment situations, they similarly face a number of potential difficulties. These include a lack of qualified teachers, as often speakers are elderly and untrained as teachers, while younger, trained teachers lack proficiency in the language. Lack of pedagogical materials and other basic resources (such as dictionaries) often hamper programs. At a more profound level, a lack of consensus about how the language is to be spoken and how it is to be written, with disagreements frequently centering around issues of orthography and codification, who has the right to teach it, have put an end to many potential programs. Access to the media is often difficult to impossible for endangered languages.

Legislation at local and national levels can either aid or impede endangered-language education. Policies which promote multilingualism and allocate resources to support educational programs can have favorable effects, while more monolingual policies which require testing or use of a single national language can have a negative impact which is difficult for a local community to overcome.

See also: Endangered Languages; Identity: Second Language; Language Teaching Traditions: Second Language; Minority Languages: Education; Minority Languages: Oppression; Second Language Attrition.

Bibliography

Fishman J (ed.) (2001). *Can threatened languages be saved?* (2nd edn.). Clevedon, England: Multilingual Matters.

Fishman J (1991). *Reversing language shift: theoretical and empirical foundations of language assistance in minority sociolinguistic perspective.* Clevedon, England: Multilingual Matters.

Hinton L (2002). *How to keep your language alive: a commonsense approach to one-on-one language learning.* Berkeley, CA: Heyday Press.

Hinton L & Hale K (eds.) (2001). *The green book of language revitalization in practice.* San Diego: Academic Press.

UNESCO (2003). 'Language vitality and endangerment: by way of introduction.' Paris: UNESCO.

Relevant Websites

http://portal.unesco.org.

Language Education for Language Revival

M Ó Laoire, The Institute of Technology, Tralee, Ireland

Language revitalization is the targeted goal of language revival efforts. It involves a reversal of language shift where people start using a language that has been moribund or threatened by extinction, so that its vitality is gradually restored (Spolsky, 1996: 6). There is a range of contexts in which language revitalization can be studied, where conditions vary considerably, e.g., contexts of nation-states, indigenous linguistic minorities in nation-states, indigenous groups in post-colonial countries, and immigrant language groups. In all these cases, however, language revitalization involves the following defining characteristics or hallmarks:

1. Adding new sets of speakers to the language crucially involving the home domain and intergenerational transmission (Spolsky, 1996; King, 2001).
2. Adding new functions by introducing the language into new domains, where it was previously unused or relatively underused.
3. The revalorization of the language to be revived (henceforth L-[r]) by the L-[r] speakers and neo-speakers (Huss *et al.*, 2003).
4. Involvement and activity on behalf of the individual and community of L-[r] speakers, and awareness that positive attitudes, action, commitment, strong acts of will, and sacrifice may be necessary to save and revitalize.

The language must gain access to domains related to socioeconomic advancement. Education is a critical domain in this respect. Huss *et al.* (2003: 4) point out that schools as well as preschools have shown themselves to be critical contexts for both language loss and revitalization. In communities where formal schooling is linked to economic advancement, without a school system in which the minority language has its proper place, all other revival efforts are likely to falter.

Whereas debate on education for language revival often centers on educational equity questions, status and institutionalization issues, and on the assessment of efficacy of varying bilingual/immersion programs, the discussion here will review specific pedagogical issues that are seen to fuel and foster language learning for language use in the speech community.

Schools and Language Revival Efforts

The debate on the role of the school in language revitalization has centered typically on schools as agents of language revival, examining the concept of language planning and language education policy (Spolsky and Shohamy, 2000) and discussing the potential of schools in the community or in national efforts to contribute to language knowledge and language use.

Fishman (1991) claims that schools have limited value in language revival in that restoration and successful survival of a threatened language essentially requires reinstating and relocating the L-[r] primarily in the home domain in parent-child transmission. Unless schools directly feed into and facilitate the reinstatement of home and family transmission, then they will always occupy a secondary role in language restoration. This does not always happen, however. It is a feature of many language revitalization movements that they overlook the crucial stage of family transmission (Fishman, 1991, Stage 6 in the Graded Intergenerational Disruption Scale) in an effort to move with undue alacrity to minority language education (May, 2001: 142).

Mindful of the shortcomings of school-based language revitalization efforts, Hornberger and King (1996: 438–439), maintain, however, that school initiatives in some contexts may promote the instruction and use of unified native languages and standardized native language literacies as well as facilitate the very kernel of the spirit of language revival. Of course, schools are also the central arena for the promotion of prescriptive norms.

Schools on their own, therefore, may be ineffective in saving threatened languages (May, 2000). Links with the speech community are critical. McCarthy (1998), for example, argues that schools must adopt a prominent position in language revitalization and maintenance efforts because schools have had destructive effects on indigenous languages in the past. Education is also the site where larger political, social, and ideological values are transmitted and reflected, the very values that fuel language revival struggle. Schools can thus become awareness-raising agents, sensitizing students to language use or lack of language use in community domains and influencing linguistic beliefs, practices, and management of the language community. Skutnabb-Kangas (2000: 570) refers to the potential of the school in this context as an agent of change. The school may also be one of the chief agents of legitimization and institutionalization

in the public domain of the L-[r] language, a counterforce of language discrimination accruing after centuries of proscription, derogation, and neglect.

Differing Contexts

Schools also operate in global and as well as in national and local contexts, involving varying standards and norms, language attitudes, multilingualism, and language prestige. Decisions on the type of program (schoolwide/immersion or targeted/one-way or two-way) (Hornberger, 1996) are best made through a process involving not just the teachers and curriculum planners, but critically involving the speech community itself. Immersion/content and language integrated learning (CLIL) and bilingual programs are often the most favored in revival contexts (e.g., Welsh: Jones, 1998; Baker, 2001; Maori: Spolsky, 1996; Benton and Benton, 2001; Irish: Coady and Ó Laoire; Slovinian in Austria: Busch, 2001; Sorbian in Germany: Elle, 2003, etc.). It cannot be suggested, however, that bilingual or immersion programs are applicable in multilingual or multicultural societies (Brann, 1981; Choudry, 2001; Benson, 2003).

Pedagogy for Language Revitalization

Despite the different contexts, common questions arise about the teaching of language in revitalization contexts. These questions focus on the optimal context and conditions for effecting revitalization, involving the type of school program, the curriculum and classroom language use and activities, the space and relationship between the L-[r] and other languages, language materials, and teacher education. Given the variety and complexity of possible contexts around the world, education policy makers and teachers need to construct the best answer to these questions in their own local contexts.

King (2001) points to one oversight in the general debate on language education for language revival: It tends to exclude reference to the specific pedagogical activities that take place in language programs. Mindful of this oversight, the remainder of this essay examines curricular and pedagogical possibilities in language revitalization programs.

Curriculum and Language Revival

One is inclined to speak about language learning programs in language revitalization contexts as if such programs were homogeneous and prone to significant variations and rates of success because of the models of teaching and learning that they comprise. King (2001) shows, for example, in the context of Quichua in the town of Saraguro in the southern highlands of Ecuador, that school-based language programs may be limited by their inappropriate pedagogical approaches, in that they remained conventional and ineffective and had limited impact in stimulating active use of the language among learners.

The contribution of sociolinguistics to language pedagogy was apparent during the last two decades of the 20th century as the focus of instruction in many L2 + FL programs broadened to include communicative competence. Communicative competence, defined as what a speaker needs to know to communicate appropriately within a particular speech community, (Saville-Troike, 1989) emerged as a basic tenet in the context of sociolinguistics and was subsequently adopted by syllabus designers and specialists in second language (L2) and foreign language (FL) instruction.

Communicative Language Teaching and L-[r]

In the final decades of the 20th century, therefore, communication in its manifold aspects has been accepted as the central tenet in language teaching (Trim, 1992).

Communicative competence and proficiency have thus become the central aim of many school-based language programs with pedagogies and language learning materials generally reflecting this central thrust of emphasis, or with the general features of communicative language teaching (CLT) being gradually absorbed into alternative or traditional methodologies. CLT, however, has particular challenges and advantages within the context of L-[r] revivalist efforts.

One challenge often bedeviling L-[r] pedagogy is the extent of native speakers. As pointed out by Little (2003) in the case of learning Irish, the number of learners of the L-[r] at any time may exceed the number of native or accustomed speakers. Pedagogical tasks thus are constrained, given the limited scope for rehearsal of interactions with native speakers. Of central concern here is the urgency or need to communicative in the L-[r] in any domain when the dominant language is perceived to be the most appropriate.

This distribution of L-[r] speaking networks may often pose, therefore, a serious problem for the learner, particularly within a communicative framework, where the relevance is wholly identified with societal use. For many schools, there is no readily identifiable speech community where such communication might be meaningful other than in communicational

transactions in the speech community. The communicative-type syllabi imply that learners, who have little or no prospect of eventually integrating into or enacting with the speech community, are asked to suspend disbelief and rehearse communicative situations, which can only be authentic or valid within the native speaker community networks inside or outside the speech community.

Second, the contraction of social contexts and domains in which the language is naturally used narrows the range of appropriate authentic materials and text-types. The absence of appropriate classroom materials is often a major obstacle.

Publishers see a limited market for developing language learning materials and resources in the L-[r], and thus teachers are forced to construct their own materials, using photocopies, drawings, and cut-out magazine clippings with superimposed handwritten captions. In doing so, they vie for students' attention and respect, especially for students who are familiar with glossy sophisticated textbooks and the high-tech learning materials of the dominant language.

Efforts in L-[r] language classrooms intent on simulating the tourist-type situations so central to communicative pedagogy of more widely used languages have worn thin with many learners (Ó Laoire, 2004). Students quickly see through the ruse and efforts to engage learners' motivation in mimicry of communicative situations, such as booking a hostel in the L-[r] in an English-speaking city, for example, or asking directions while working from a map, and have been doomed to arouse at best a benign indifference, even apathy, among the most eager learners.

However, the shift to communicative language teaching has been useful in language instruction for promoting language use in the speech community. This has positive underlying potential to foster active language use of the L-[r], in that it draws deliberate attention to domains where the language is used meaningfully in transactions and at the same time introduces the learners to new situations where the language potentially could be used. Thus, in developing receptive competences, for example, learners hear the language as it used in the speech community, or read authentic texts where printed and literacy texts are available. In developing productive competences, learners engage in role-plays to prepare for interactions with native speakers, neo-speakers, or accustomed speakers, even if such interactions might never occur. This approach is somewhat at variance with the remarkably unprogressive methods and passive activities described by King (2003).

It is not always easy, even for committed students, to communicate, or even to know how and when to communicate, with native/neo-speakers speakers of the L-[r], either inside or/and outside the speech community. Learners at this crucial integration-threshold stage often think that their command of the language is not good enough and compare their own efforts unfavorably with the standard of the target network group. Unfortunately, such learners often give up. This points to a need not only for more research into the sociolinguistic and motivational variables of integration, but also for preliminary studies of inter-language pragmatics in the case of L-[r] speakers.

Neighborhood domains and home domains in the early stages of language revival are often insufficient for the learner to sustain or to develop proficiency through use. The school alone may indeed be the only source of language learning and inter-language development, and they may never be reinforced by integration into the speech community. This is, perhaps, the greatest challenge to acquiring the L-[r] and is a strong argument for immersion or content-based language learning as being the best suited pedagogically in the context of revivalist efforts.

Different Expectations from Pedagogy

Schools on their own will not change language behavior. A complication often is the different goals, definitions, and measurements of success employed by different revivalist groups for whom aims tend to vary considerably:

1. One group may aim only for the language being taught in the schools (one type of syllabus) – once it is being taught, they are happy.
2. Another group might aim to reintroduce the language into families and thereby secure intergenerational transmission.

These two approaches would necessitate different syllabuses. The former would aim to teach the language in some de-contextualized context or for cultural reasons, whereas the latter would stress learning the language for active use and for use in the micro-domain of the family itself. This forces us to answer the question: Why is L-[r] being taught in the first place?

1. For short-term motivational fulfillment, i.e., communicating in the classroom?
2. To secure in the longer term the use of the language in the family domain and thus secure intergenerational transmission.

Both aims can coalesce pedagogically, of course. There is a difference, nonetheless, between language learning and language acquisition. Language that is

learned may be forgotten, if it is not retrieved from short-term memory, used, or activated on a regular basis. Language that is acquired is automated or automatically available. It is possible to extend and drive inter-language and fuel acquisition in the classroom if the language is used meaningfully. The communicative and purposeful use of language offers this opportunity. We need not only to give our learners practice in the language through scaffolding, but we must also create opportunities for our learners to talk meaningfully and therein process language. The more exposure to the L-[r] as in immersion/submersion CLIL programs and the more opportunities to use the language in a meaningful way, the better the chances of acquisition.

Final Considerations: Language Revitalization and Instructed Language Acquisition

This discussion suggests that real-life use of language should be reflected to a large extent in the texts and activities of the classroom. As these activities do not depend on immediate reinforcement outside the classroom, pedagogical activities can be seen as an end in themselves and reflect real peer-to-peer or teacher-to-student communication. The classroom, therefore, becomes a valid communicative situation, exploitable to a large extent as a valuable resource for instructed language acquisition (Ellis, 1990). Critical to construction of contexts of acquisition in the classroom is the concern with learning rather than teaching (Little *et al.*, 2000). In the case of acquisition of L-[r] in instructed contexts, much remains to be investigated. It is likely, however, that research in this will conform to a significant extent to the following basic findings of L2 studies to date.

It can be expected, therefore, that significant amounts of language exposure are required. Teaching the L-[r] as a subject only is unlikely to succeed. Experience from immersion programs indicates that a greater investment of hours of exposure may be necessary to achieve acquisition (King, 2001: 216). Second, learners must engage in meaningful interactions in the classroom wherein negotiation for meaning triggers interactional adjustment by native speakers (Long, 1996). The processes of acquisition as well as being contingent on modification of native speaker and comprehensible input depend equally on meaningful interaction with the target language. Finally, as well as teaching and exposing our learners to the L-[r], as teachers we need to equip them with skills in how to learn the language (and other languages) and how to seek out opportunities to use it outside the classroom. Such an approach can imbue teachers and learners alike with a new creative enthusiasm for language in general and create a language awareness that is facilitative of acquisition.

See also: Language Education for Endangered Languages.

Bibliography

Baker C (2001). *Foundations of bilingual education and bilingualism* (3rd edn.). Clevedon: Multilingual Matters.

Benson C (2003). 'Possibilities for educational language choice in multilingual Guinea-Bissau.' In Huss L *et al.* (eds.). 67–88.

Benton R & Benton N (2001). 'RSL in Aotearoa/New Zealand 1989–1999.' In Fishman J A (ed.) *Can threatened languages be saved?* Clevedon: Multilingual Matters. 423–450.

Brann C (1981). *Trilingualism in language planning for education in Sub-Saharan Africa.* Paris: UNESCO.

Busch B (2001). 'Slovenian in Carinthia.' In Extra G & Gorter D (eds.). 119–136.

Cenoz J (2001). 'Basque in Spain and France.' In Extra G & Gorter D (eds.). 45–58.

Choudry A (2001). 'Linguistic minorities in India.' In Extra G & Gorter D (eds.). 391–406.

Coady M & Ó Laoire M (2002). 'Mismatches in language policy and practice in education.' *Language Policy 1(2),* 143–158.

Elle L (2003). 'Revitalisation through immersion – A new road to Sorbian language maintenance.' In Huss L *et al.* (eds.). 89–102.

Ellis R (1990). *Instructed second language acquisition.* Oxford: Blackwell.

Fishman J A (1991). *Reversing language shift. Theoretical and empirical foundations of assistance to threatened languages.* Clevedon: Multilingual Matters.

Extra G & Gorter D (eds.) (2001). *The other languages of Europe.* Clevedon: Multilingual Matters. 119–136.

Hornberger N (1996). 'Language in education.' In McKay S L & Hornberger N (eds.) *Sociolinguistics and language teaching.* Cambridge: Cambridge University Press. 449–473.

Hornberger N & King K A (1996). 'Language revitalization in the Andes: Can the schools reverse language shift?' *Journal of Multilingual and Multicultural Development 17(6),* 427–441.

Huss L, Camilleri Grima A & King K A (eds.) (2003). *Transcending monolingualism. Linguistic revitalization in education.* Lisse: Swets & Zeitlinger.

Jones M (1998). *Language obsolescence and revitalization.* Oxford: Clarendon Press.

King K A (2001). *Language revitalization. Processes and prospects.* Clevedon: Multilingual Matters.

King K A (2003). 'Language pedagogy and language revitalization: Experiences from Ecuadorian Andes and beyond.' In Huss L *et al.* (eds.). 89–102.

Lee McKay S & Hornberger N (eds.) (1996). x*Sociolinguistics and Language Teaching*. Cambridge University Press.
Little D (2003). *Languages in the post-primary curriculum: A discussion paper.* Dublin: NCCA. http://www.ncca.ie.
Little D, Dam L & Timmer J (eds.) (2000). *Focus on learning rather than teaching: Why and how?* Dublin: CLCS, Trinity College.
Long M (1996). 'The role of linguistic environment in second language acquisition.' In Richie W C & Bhatia T K (eds.) *Handbook of second language acquisition.* Rowley, MA: Newbury House. 377–393.
May S (2001). *Language and minority rights. Ethnicity nationalism and the politics of language.* Harlow: Longman Pearson Education.
McCarthy T L (1998). 'Schooling, resistance and American Indian languages.' *International Journal of the Sociology of Language 132,* 27–41.
Ó Laoire M (2004). 'Bilingualism in school settings in Ireland: Perspectives on Irish in the L2 curriculum.' In Cohen J, McAlister K, Rolstad K & MacSwan J (eds.) *ISB4: Proceedings of the 4th International Symposium on Bilingualism.* Somerville, MA: Cascadilla Press.
Saville-Troike M (1989). *The ethnography of communication.* Oxford: Basil Blackwell.
Skutnabb-Kangas T (2000). *Linguistic genocide in education or worldwide diversity and human rights?* New Jersey: Lawrence Erlbaum.
Spolsky B (1989). 'Maori bilingual education and language revitalization.' *Journal of Multilingual and Multicultural Development 10(2),* 89–106.
Spolsky B (1996). 'Conditions for language revitalization: A comparison of the cases of Hebrew and Maori.' In Wright S (ed.) *Language and the state.* Clevedon: Multilingual Matters. 5–29.
Spolsky B (2003). *Language policy.* Cambridge University Press.
Spolsky B & Shohamy E (2000). 'Language practice, language ideology, and language policy.' In Lambert R D & Shohamy E (eds.) *Language policy and pedagogy.* Philadelphia/Amsterdam: John Benjamins. 1–42.
Trim J (1992). 'Language teaching in the perspective of the predictable requirements of the twenty-first century.' *AILA Review 9,* 7–21.
Williams C (2001). 'Welsh in Great Britain.' In Extra G & Gorter D (eds.). 59–82.

Language Education of the Deaf

J Kyle, University of Bristol, Bristol, UK

Deafness has always existed, and for the last 240 years, it has merited the attention of formal education. According to Samuel Johnson in 1775, the problems were already solved.

> ... in Edinburgh ... a college of the deaf and dumb, who are taught to speak .... The improvement of Mr. Braidwood's pupils is wonderful. They not only speak, write, and understand what is written but if he that speaks looks towards them and modifies his organs by distinct and full utterance, they know so well what is spoken that ... they hear with the eye (Johnson, 1775: 143–144).

In that statement, we have the whole of the mystery that has pervaded deaf education since Mr. Johnson's visit to Scotland. How is it that some deaf children appear to learn to speak, lip-read, read, and write, and the remainder, the majority, have stubbornly refused to respond to the efforts of educators? It was, and remains, a matter of schism, doctrine, and even deception concerning the methods and outcomes of the process through which deaf children pass.

One variable in the equation is hearing loss. It was not until the 1920s that hearing loss was reliably measured. Prior to that, deaf children were mixed with partially hearing children and with children who became deaf after the age when they had learned to speak. It was relatively easy to convince benefactors that a child's competence was the result of treatment in school.

A second variable is the nature of the schooling. At first, deaf children boarded at school, and later in cities they could be day pupils. By the late 1940s, there were special units set up in mainstream schools to cater to those children who were partially hearing. Latterly, on the bandwagon of civil rights and a general philosophy of integration, deaf children were placed directly in mainstream schools, sometimes alone, but supported by the provision of aids to hearing and by a visiting teacher.

A third variable is the use of hearing aids. Although widely available since the 1950s in most developed countries, they are relatively rare in developing countries. By early detection and application of hearing aids, many children deemed deaf in the past are now seen as partially hearing and are able to interact with hearing peers in a mainstream school. A more sophisticated and more effective version of hearing aid is the cochlear implant, an operation to implant a device in the inner ear to restore hearing.

A fourth variable, though insufficiently understood, is parental involvement. As with all human special

Lee McKay S & Hornberger N (eds.) (1996). x*Sociolinguistics and Language Teaching*. Cambridge University Press.
Little D (2003). *Languages in the post-primary curriculum: A discussion paper.* Dublin: NCCA. http://www.ncca.ie.
Little D, Dam L & Timmer J (eds.) (2000). *Focus on learning rather than teaching: Why and how?* Dublin: CLCS, Trinity College.
Long M (1996). 'The role of linguistic environment in second language acquisition.' In Richie W C & Bhatia T K (eds.) *Handbook of second language acquisition.* Rowley, MA: Newbury House. 377–393.
May S (2001). *Language and minority rights. Ethnicity nationalism and the politics of language.* Harlow: Longman Pearson Education.
McCarthy T L (1998). 'Schooling, resistance and American Indian languages.' *International Journal of the Sociology of Language 132,* 27–41.
Ó Laoire M (2004). 'Bilingualism in school settings in Ireland: Perspectives on Irish in the L2 curriculum.' In Cohen J, McAlister K, Rolstad K & MacSwan J (eds.) *ISB4: Proceedings of the 4th International Symposium on Bilingualism.* Somerville, MA: Cascadilla Press.
Saville-Troike M (1989). *The ethnography of communication.* Oxford: Basil Blackwell.
Skutnabb-Kangas T (2000). *Linguistic genocide in education or worldwide diversity and human rights?* New Jersey: Lawrence Erlbaum.
Spolsky B (1989). 'Maori bilingual education and language revitalization.' *Journal of Multilingual and Multicultural Development 10(2),* 89–106.
Spolsky B (1996). 'Conditions for language revitalization: A comparison of the cases of Hebrew and Maori.' In Wright S (ed.) *Language and the state.* Clevedon: Multilingual Matters. 5–29.
Spolsky B (2003). *Language policy.* Cambridge University Press.
Spolsky B & Shohamy E (2000). 'Language practice, language ideology, and language policy.' In Lambert R D & Shohamy E (eds.) *Language policy and pedagogy.* Philadelphia/Amsterdam: John Benjamins. 1–42.
Trim J (1992). 'Language teaching in the perspective of the predictable requirements of the twenty-first century.' *AILA Review 9,* 7–21.
Williams C (2001). 'Welsh in Great Britain.' In Extra G & Gorter D (eds.). 59–82.

Language Education of the Deaf

J Kyle, University of Bristol, Bristol, UK

Deafness has always existed, and for the last 240 years, it has merited the attention of formal education. According to Samuel Johnson in 1775, the problems were already solved.

> ... in Edinburgh ... a college of the deaf and dumb, who are taught to speak .... The improvement of Mr. Braidwood's pupils is wonderful. They not only speak, write, and understand what is written but if he that speaks looks towards them and modifies his organs by distinct and full utterance, they know so well what is spoken that ... they hear with the eye (Johnson, 1775: 143–144).

In that statement, we have the whole of the mystery that has pervaded deaf education since Mr. Johnson's visit to Scotland. How is it that some deaf children appear to learn to speak, lip-read, read, and write, and the remainder, the majority, have stubbornly refused to respond to the efforts of educators? It was, and remains, a matter of schism, doctrine, and even deception concerning the methods and outcomes of the process through which deaf children pass.

One variable in the equation is hearing loss. It was not until the 1920s that hearing loss was reliably measured. Prior to that, deaf children were mixed with partially hearing children and with children who became deaf after the age when they had learned to speak. It was relatively easy to convince benefactors that a child's competence was the result of treatment in school.

A second variable is the nature of the schooling. At first, deaf children boarded at school, and later in cities they could be day pupils. By the late 1940s, there were special units set up in mainstream schools to cater to those children who were partially hearing. Latterly, on the bandwagon of civil rights and a general philosophy of integration, deaf children were placed directly in mainstream schools, sometimes alone, but supported by the provision of aids to hearing and by a visiting teacher.

A third variable is the use of hearing aids. Although widely available since the 1950s in most developed countries, they are relatively rare in developing countries. By early detection and application of hearing aids, many children deemed deaf in the past are now seen as partially hearing and are able to interact with hearing peers in a mainstream school. A more sophisticated and more effective version of hearing aid is the cochlear implant, an operation to implant a device in the inner ear to restore hearing.

A fourth variable, though insufficiently understood, is parental involvement. As with all human special

needs areas, the extent to which the parents show an interest, offer home tuition, and are in a position financially to support the child's special needs, tends to create greater or lesser level of achievement.

In all cases, claims are made that attention to these variables will eradicate the problem of deafness. This has gradually narrowed the problem from being defined as comprising all those with a hearing loss to those who are deaf from infancy, in the provision of special education. The deaf population has 'changed' since 1775.

Methods

In the absence of medical or technical solutions, education has sought remedies of its own. These have succeeded in polarizing the field for more than 200 years and in leaving deaf children with poor levels of achievement and language performance (Binet, 1910; Conrad, 1979; Powers *et al.,* 1998).

The German Method, in evidence in the late 18th century, emphasized the teaching of speech to the exclusion of gesture or signing. Teachers spoke to the deaf children and forbade them to sign. It became the oral method which dominated deaf education for nearly 200 years.

The French Method adapted the natural signing of deaf people to 'methodize' the sentence construction so that it might better reflect written and spoken French. Deaf children were allowed to sign to each other (something forbidden by the oral method), and their natural language communication was valued. American educational practice was greatly influenced by the French Method, and this was later to evolve into Total Communication (Denton, 1976) – a formal use of speaking, signing, and signed speech.

The British (or combined) Method, which Johnson saw, was primarily oral but used finger-spelling and allowed the children to sign out of class. By the late 19th century, this method had all but disappeared, as deaf teachers were removed by the change toward professionalization of teaching and the rise to dominance of a medically inspired, teacher training program in Manchester. For most of the 20th century, Europe used the oral method.

Natural Language of Deaf Children

However, in the 1970s, on the basis of linguistic research in the United States and Europe, sign *language* began to be described, and deaf culture began to be promoted by deaf people, by professionals, and then by hearing parents. Teachers began to teach through sign language and to create a bilingual framework.

First noted as early as 1644 in England, signed languages have now been identified and described in nearly every country in the world (Klima and Bellugi, 1979; Kyle and Woll, 1985; Schulmeister and Reinitzer, 2002 – *see* **Sign Languages of the world** (00243)). Sign languages are learned naturally in families with deaf parents (about 6% of all cases of childhood deafness). What became apparent was their complexity, flexibility, their ease of acquisition among deaf children in deaf families, and their potential for deployment in family intervention and education (Johnson *et al.,* 1989).

The Situation Today

Oral Approaches

There continues to be considerable debate on the methods for teaching deaf children. Surprisingly, perhaps, given the extent of problems reported by deaf people themselves (Lane *et al.,* 1996), oral education continues to be applied in many schools in most parts of the world. This is a function of the naturalness of native language use by the hearing educators and parents, and of the priority to normalize the child. The method has evolved from the former speech training exercises, drills, and compulsion to talk, toward a natural or aural method where emphasis is placed on hearing and communication. Watson (1998) says "... it is assumed that deaf children can utilise their hearing via hearing aids in order to develop the sounds of speech" (Watson, 1998: 72).

The system relies on early detection of hearing loss (preferably at birth) and the immediate provision of hearing aids, and then the monitoring of their use. It requires preschool programs and support systems usually found only in the developed world. It is said to require a considerable commitment from parents who have the primary responsibility to continue to provide speech experiences to the child even in the face of slow development (which is considered the norm). The advantage of the approach is that it allows a denial of deafness and assertion of the child's normality – because the primary aim is to make the child function in speech, reading, and writing as a hearing person – as Johnson originally described.

Total Communication (TC)

Typically this approach involves teachers who speak and sign simultaneously (Baker and Knight, 1998), even though the purists would argue that sign language (with its own grammatical structure) should be used alongside speaking and writing). Initially, this required the construction of artificial signs in order to

allow the precise mapping of speech – i.e., sign language was thought to be limited in vocabulary. This remains a concern for hearing educators, and often there are committees set up to devise new signs for use in TC programs. Simultaneous communication requires that the teacher is able to sign all elements of words as they are spoken. Not only is this considered to affect the signing (i.e., to make it unintelligible), it is also claimed to affect the speech production of the teacher (Hyde *et al.*, 1998).

Because of the dominance of the oral approach, often deaf children arrive in TC programs later than is ideal and commonly as oral failures (this may even be claimed as a natural application of the oral approach whereby, if the children do not succeed, they still have the choice of a signing program later (Watson, 1998)). Teachers have to learn the signed *method* (although this often does not require learning of the signed *language* – only the signed speech variant).

For these reasons, TC programs in developed countries tend to have shifted toward bilingualism, which is now a major development area in deaf education worldwide.

Sign Bilingualism

Instead of concentrating solely on the child's spoken language development, bilingual approaches emphasize the early mastery of sign language, while ensuring that the child has access to the spoken, heard, and written language of the society. This usually involves parents learning to sign, the presence of deaf role models, and contact with the deaf community (Young, 1999). At school, it involves policies to ensure equality of status of the two languages, curriculum access through and assessment in sign language as well as the written language, and staff training in sign language (Pickersgill, 1998). The underpinning of the approach is the extensive literature which supports bilingual education worldwide (Cummins, 2000).

For these programs to be effective, there has to be a shift in perception of the child from 'hearing-impaired' to 'culturally deaf' and thereby different from the parents and other family members. Advanced programs exist in Sweden (Ahlgren, 1990), in Denmark (Hansen, 1990), and there are descriptions of programs in Italy (Volterra, 1990), in Holland (Knoors, 1995), in the United States (Nover and Andrews, 1998), in Russia (Komarova, 1998), in Latin America (Skliar, 1997), and in China (Calloway, 2000).

Advantages in the programs are seen to be that they offer a major break from the approaches used in deaf education for 200 years and for the first time emphasize sign language information and the cultural experience of the child, that they provide a natural language framework, and that they are theoretically sound being based on extensive work on bilingual education in the spoken language field.

Do the Methods Work?

Most proponents of these approaches tend to shift uncomfortably in the glare of research analysis, and in reality it is difficult to answer the question of effectiveness directly. A thorough review of deaf achievement by Powers, Gregory, and Thoutenhoofd (1998) found flaws in nearly all of the published research studies in this area. As with all educational research, the true results cannot be determined until many years after program implementation. Bilingual approaches are in the early stages of development. What is significant, however, is that for the first time, the deaf community is involved and deaf people are employed in preschool programs and in schools. Their native experience of learning and their knowledge of the visual world are extremely important in the education of the deaf child.

Typically, national examinations of deaf children at the end of their schooling show major problems in achievements in speech, lip-reading, and reading. However, even in oral programs, deaf children become fluent in sign language (even where there have been measures taken to prevent this). It would seem obvious that because education focuses on learning, the most effective language (i.e., sign language) should be used as a primary tool and that mastery of signed and then, spoken/written languages should be the priority.

See also: Bilingual Education; Sign Language: Interpreting; Sign Languages of the World.

Bibliography

Ahlgren I (1990). 'Swedish conditions: sign language in deaf education.' In Prillwitz S & Vollhaber T (eds.).

Baker R & Knight P (1998). 'Total communication – current policy and practice.' In Gregory *et al.* (eds.).

Binet A (1910). 'An investigation concerning the value of the oral method.' *American Annals of the Deaf 55,* 4–33.

Calloway A (2000). *Deaf children in China.* Washington, D.C.: Gallaudet University Press.

Conrad R (1979). *The deaf school child.* London: Harper & Row.

Cummins J (2000). *Language power and pedagogy.* Clevedon, UK: Multilingual Matters.

Denton D (1976). *The philosophy of total communication, supplement to the British Deaf News.* Carlisle: British Deaf Association.

Gregory S, Knight P, McCracken W, Powers W & Watson L (eds.) (1998). *Issues in deaf education.* London: David Fulton.
Hansen B (1990). 'Trends in the progress towards bilingual education for deaf children in Denmark.' In Prillwitz S & Vollhaber T (eds.).
Hyde M, Power D & Leigh G (1998). 'Oral-only and simultaneous communication speech characteristics of teachers of the deaf.' In Weisel A (ed.) *Issues unresolved: new perspectives on language and deaf education.* Washington, DC: Gallaudet University Press.
Johnson R, Liddell S & Erting C (1989). *Unlocking the curriculum.* Washington, DC: Gallaudet Research Institute.
Johnson S (1775, reprinted 1996). *Journey to the Hebrides.* Edinburgh: Canongate Classics.
Klima E S & Bellugi U (1979). *The signs of language.* Cambridge, Mass: Harvard University Press.
Knoors H (1995). *The inevitable bilingualism of deaf children. Tel Aviv International Congress on the Education of the Deaf.*
Komarova A (1998). 'Teaching English to deaf children in a Moscow bilingual school.' In Zaitseva G, Komarova A & Pursglove M (eds.) *Deaf children and bilingual education.* Moscow: Bilingual School.
Kyle J G & Woll B (1985). *Sign language.* Cambridge: Cambridge University Press.
Lane H, Hoffmeister R & Bahan B (1996). *A journey into the deaf world.* San Diego, CA: Dawn Sign.
Nover S M & Andrews J F (1998). *Critical pedagogy in deaf education: deaf methodology and staff development.* Santa Fe: New Mexico School for the Deaf.
Pickersgill M (1998). 'Bilingualism – current policy and practice.' In Gregory *et al.* (eds.).
Powers S, Gregory S & Thoutenhoofd E D (1998). *The educational achievements of deaf children, RR65.* London: DfEE.
Prillwitz S & Vollhaber T (eds.) (1990). *Sign language research and application.* Hamburg: Signum.
Schulmeister R & Reinitzer H (2002). *Progress in sign language research.* Hamburg: Signum.
Skliar C (1997). *La educacíon de los sordos.* Mendoza: Universidad Nacional de Cuyo.
Voterra V (1990). 'Sign language acquisition and bilingualism.' In Prillwitz S & Vollhaber T (eds.).
Watson L (1998). 'Oralism: current policy and practice.' In Gregory S *et al.* (eds.).
Young A (1999). 'Hearing parents adjustment to a deaf child.' *Journal of Social Work Practice 13,* 157–176.

Language Education Policies in Africa

K Heugh, Human Sciences Research Council, Cape Town, South Africa

Historical Background to Language Education Policy

Language education policies in Africa, since the 19th-century partition of the continent by Europe, have been out of kilter with the realities of its speech communities. These policies are beset with contradictory, inappropriate implementation strategies and generally give priority to former colonial languages that, after 120 years of use in formal education systems, remain little known and seldom used by approximately 80% of the people. The 10–20% (Heine, 1992) of African people who do have high levels of proficiency in one of these languages are those who occupy the upper echelons of political power, the economy, and the education systems, and they live in urban areas.

The African continent is, at the same time, richly endowed with linguistic diversity, education systems that began thousands of years ago, and numerous written scripts (e.g., Battestini, 1997), as well as many languages that have not yet been reduced to writing systems. Historiographic and contemporary accounts of the continent usually emphasize its linguistic complexity. Somewhere between 1250 and 2100 (depending on identification criteria) or 30% of the world's languages are found in Africa (Ouane, 2003).

However, most literature focuses on an apparent dearth of written languages and on the contributions of Christian missionaries and colonial administrations to literacy, writing, and education systems. The significance of pre-colonial language and educational practices, recorded in languages and scripts that are little understood in the West and North, has gone largely unrecognized. This article begins with pre-colonial systems of 'language management' (Spolsky, 2004) in order to shed light on the origins of mismatched language policy decisions and implementation plans.

Richly textured oral traditions have been significant conduits for the preservation and expansion of history, literature, and knowledge systems, as well as for their transmission from one generation to the next for thousands of years. These oral traditions were not, however, the only vehicle for language and education development. The Ancient Egyptians' use of hieroglyphics influenced many writing systems in countries further south. The *Gicandi* script of the

Gregory S, Knight P, McCracken W, Powers W & Watson L (eds.) (1998). *Issues in deaf education.* London: David Fulton.
Hansen B (1990). 'Trends in the progress towards bilingual education for deaf children in Denmark.' In Prillwitz S & Vollhaber T (eds.).
Hyde M, Power D & Leigh G (1998). 'Oral-only and simultaneous communication speech characteristics of teachers of the deaf.' In Weisel A (ed.) *Issues unresolved: new perspectives on language and deaf education.* Washington, DC: Gallaudet University Press.
Johnson R, Liddell S & Erting C (1989). *Unlocking the curriculum.* Washington, DC: Gallaudet Research Institute.
Johnson S (1775, reprinted 1996). *Journey to the Hebrides.* Edinburgh: Canongate Classics.
Klima E S & Bellugi U (1979). *The signs of language.* Cambridge, Mass: Harvard University Press.
Knoors H (1995). *The inevitable bilingualism of deaf children. Tel Aviv International Congress on the Education of the Deaf.*
Komarova A (1998). 'Teaching English to deaf children in a Moscow bilingual school.' In Zaitseva G, Komarova A & Pursglove M (eds.) *Deaf children and bilingual education.* Moscow: Bilingual School.
Kyle J G & Woll B (1985). *Sign language.* Cambridge: Cambridge University Press.
Lane H, Hoffmeister R & Bahan B (1996). *A journey into the deaf world.* San Diego, CA: Dawn Sign.
Nover S M & Andrews J F (1998). *Critical pedagogy in deaf education: deaf methodology and staff development.* Santa Fe: New Mexico School for the Deaf.
Pickersgill M (1998). 'Bilingualism – current policy and practice.' In Gregory *et al.* (eds.).
Powers S, Gregory S & Thoutenhoofd E D (1998). *The educational achievements of deaf children, RR65.* London: DfEE.
Prillwitz S & Vollhaber T (eds.) (1990). *Sign language research and application.* Hamburg: Signum.
Schulmeister R & Reinitzer H (2002). *Progress in sign language research.* Hamburg: Signum.
Skliar C (1997). *La educacíon de los sordos.* Mendoza: Universidad Nacional de Cuyo.
Voterra V (1990). 'Sign language acquisition and bilingualism.' In Prillwitz S & Vollhaber T (eds.).
Watson L (1998). 'Oralism: current policy and practice.' In Gregory S *et al.* (eds.).
Young A (1999). 'Hearing parents adjustment to a deaf child.' *Journal of Social Work Practice 13,* 157–176.

Language Education Policies in Africa

K Heugh, Human Sciences Research Council, Cape Town, South Africa

Historical Background to Language Education Policy

Language education policies in Africa, since the 19th-century partition of the continent by Europe, have been out of kilter with the realities of its speech communities. These policies are beset with contradictory, inappropriate implementation strategies and generally give priority to former colonial languages that, after 120 years of use in formal education systems, remain little known and seldom used by approximately 80% of the people. The 10–20% (Heine, 1992) of African people who do have high levels of proficiency in one of these languages are those who occupy the upper echelons of political power, the economy, and the education systems, and they live in urban areas.

The African continent is, at the same time, richly endowed with linguistic diversity, education systems that began thousands of years ago, and numerous written scripts (e.g., Battestini, 1997), as well as many languages that have not yet been reduced to writing systems. Historiographic and contemporary accounts of the continent usually emphasize its linguistic complexity. Somewhere between 1250 and 2100 (depending on identification criteria) or 30% of the world's languages are found in Africa (Ouane, 2003).

However, most literature focuses on an apparent dearth of written languages and on the contributions of Christian missionaries and colonial administrations to literacy, writing, and education systems. The significance of pre-colonial language and educational practices, recorded in languages and scripts that are little understood in the West and North, has gone largely unrecognized. This article begins with pre-colonial systems of 'language management' (Spolsky, 2004) in order to shed light on the origins of mismatched language policy decisions and implementation plans.

Richly textured oral traditions have been significant conduits for the preservation and expansion of history, literature, and knowledge systems, as well as for their transmission from one generation to the next for thousands of years. These oral traditions were not, however, the only vehicle for language and education development. The Ancient Egyptians' use of hieroglyphics influenced many writing systems in countries further south. The *Gicandi* script of the

Kikuyu in Kenya, the *Nsibidi* pictograms of the Efik in Nigeria, the Mende script of Sierra Leone, and the Loma script of Liberia, as well as Dogon, Bambara, and Bambum (Cameroon) scripts, show similarities to Egyptian hieroglyphics (Battestini, 1997; Ki-Zerbo, 2003). Modern Ethiopian languages, such as Amharic and Tigrinya (Tigrigna), have emerged from the ancient Ethiopic language, Giiz (Ge'ez), a south Semitic language in which the Holy Book, *Kebra nagast,* is written (Bloor and Tamrat, 1996).

Religion, particularly the spread of Islam since the 7th century A.D., played a significant role in the development of written texts. It also influenced early examples of explicit language policy. Abdulaziz (2003: 183) suggested that African converts to Islam "were expected to acquire basic literacy in Arabic" in order to recite Koranic texts. A similar process occurred with Christianity in Ethiopia where the literary use of Giiz peaked between the 7th and 13th centuries (Bloor and Tamrat, 1996). The spread of Islam led to the development of *Ajami* script, which is the use of Arabic for the transcription of other languages. In West Africa, Ajami was used in the development of scripts for Fulfulde, Mende, Wolof, Hausa, and Yoruba in Burkina Faso, Mali, Nigeria, and Senegal. It has also been used for Kiswahili and Somali in East Africa (Abdulaziz, 2003; Alidou, 2004). The early 19th-century scholar, Sheik Dan Fodio, argued that Islam in Africa took on a distinctively different form from that of the Arab world (Djian, 2004). Missionary and trading activity across the Sahel (areas south of the Sahara) led to the establishment of 'Islamo-Sudanese' centers. By the 15th century, 25 000 students were attending the Sankore university-mosque of Timbuktu (Djian, 2004).

The recent discovery of 15 000 documents, dating from the 12th to the 16th centuries in Timbuktu, is reason to revise earlier histories. The documents, which address religion, philosophy, music, astronomy, history, law, medicine, mathematics, and commerce, provide extensive evidence of literacy and formal education in this area. Although mostly written in Arabic, they include several other languages, such as Hausa, Wolof, Fulani (Fulfulde), Juula (Dyula), and Tamashek (Tamajeq) (Djian, 2004). The Tuareg had developed their own script, Tifinagh, for the Tamajeq language, which was used in Mauritania, Mali, and Niger (Alidou, 2004). Dan Fodio (1754–1817) established an Islamic library in Kano in northern Nigeria and translated Greek mathematics into Hausa, but his work and library were destroyed by British military action in 1903 (Boahen, 2003; Alidou, 2004).

Communication across the continent occurred, and continues today, via a network of smaller local languages and larger, vehicular languages (Djité, 1993). There are a few countries where communication occurs predominantly through one language; for example, in Rwanda, Burundi, Swaziland, Lesotho, Somalia, and Madagascar. For the most part, however, high levels of multilingualism were and continue to be common. Fardon and Furniss (1994) argued that the Western concept of 'lingua franca' is inadequate in the African context. The African lingua franca is not "a single language but . . . a multilayered and partially connected language chain that offers a choice of varieties and registers in the speaker's immediate environment" (Fardon and Furniss, 1994: 4). Other authors (e.g., Djité, 1993) referred to the widespread use of small local languages alongside the larger regional languages of wider communication, such as Hausa, Fulfulde, Yoruba, and Kiswahili, across much of the continent. Indigenous education systems provided access to these language chains and communicative patterns. The use of two or more languages for specific functions – the small local language in immediate neighborhoods, regional lingua francas for trade, and Arabic or Giiz for religious purposes – was therefore common.

During the second half of the 19th century, shifting power relations in Europe and the impact of industrialization resulted in political aspirations toward African territory. The Berlin Conference of 1884–1885 identified 'spheres of influence' and accelerated the appropriation or conquest of Africa. A series of treaties and occupation by force resulted in a partition that had little to do with earlier African political or linguistic boundaries (see Chumbow, 1999). By 1914, only Ethiopia and Liberia remained free of European domination. Colonization was accompanied by political, social, economic, educational, and, particularly, linguistic consequences.

Although pre-colonial education was itself elitist and excluded many through its "sacred/taboo/secret character" (Djité, personal communication, 2004), the introduction of European structures of domination undermined the foundations of African education and language management. Traditional institutions of power and social structures were appropriated or replaced and thus weakened systemically (Elaigwu and Mazrui, 2003). Military invasions and attempts at resistance interrupted formal education in the cities of West and East Africa. The colonial powers neither grasped the significance and efficacy of nor saw much benefit in maintaining existing systems (Alidou, 2004). The functional use of indigenous languages and writing systems was underestimated or deliberately suppressed. Once European languages were introduced for high-status functions, especially in education, indigenous languages simultaneously

lost value and, in many instances, were rendered invisible.

Colonial education systems were designed to cater to a small elite who would "serve European economic and political interests" (Alidou, 2004: 199). The majority, who might have previously taken advantage of indigenous education, found themselves without access to either system, and thus, the continent has for the last century witnessed large-scale under-education and failure. More than half (52.7%) of Africans are illiterate, and of the little more than 50% of school-aged children enrolled in school, the majority are unlikely to reach secondary school (Bamgbose, 2000). A significant reason for this high dropout rate is the emphasis on the educational use of languages that most children do not understand: "(M)ost African children merely pass through the school without the school passing through them" (Obanya, 1999: 123).

Colonial Period: Contradictory and Parallel Processes

The partition of Africa brought specific changes to the linguistic ecology of the continent. The French colonized most of West Africa, Algeria, Tunisia, Morocco, and Madagascar. The British took control of most of East Africa and a large portion of Southern Africa and Nigeria. Angola, Mozambique, Guinea Bissau, São Tome and Principe, and the Cape Verde islands went to Portugal, whereas Spain occupied a small part of Morocco, Western Sahara, and Equatorial Guinea. The Belgians retained control over Belgian Congo (Zaire, later the Democratic Republic of Congo), and the Italians initially managed to secure part of Somalia, Eritrea, and eventually Libya. They also took Ethiopia by military conquest for short periods between 1889–1896 and 1936–1941, after which they were displaced by Britain. The Germans occupied and controlled the areas now known as Tanzania, Cameroon, Togo, Rwanda, Burundi, and Namibia until World War I; after that, their territory was redivided among Britain, France, Belgium, and South Africa.

The Europeans brought contradictory processes affecting language policy and management. First, there was an attempt to imprint the European concept of the nation-state with a single, national language onto each African country (Bamgbose, 1994). This effort effectively rendered local languages politically invisible and fostered assimilation into the colonial language and power structures for the small (new) elite. Second, an apparent multiplication or division of languages accompanied colonization. The arbitrary partition of Africa ignored existing linguistic boundaries and cut across "linguistic frontiers" (Fardon and Furniss, 1994), thus dividing speech communities (see also Chumbow, 1999). Linguistic fragmentation also resulted from the work of missionaries who, often without sufficient linguistic expertise, produced different orthographies for closely related language varieties. They thus increased the distance among languages and possibly 'invented' languages (Makoni, 1998; Ranger, 1989). Although by 1906 missionaries in Southern Africa sought to reverse this process, it was too late (Heugh, 2003).

A third, dynamic process emerged from among the speech communities themselves in the form of pidgins (e.g., in Cameroon and Nigeria) and of Creoles, especially along the Atlantic coastal areas. The Creoles of Gambia, Cameroon, and Sierra Leone borrowed terms from English, whereas the *crioulo* of Cape Verde, São Tomé and Principe, and Guinea Bissau borrowed from Portuguese (Sow and Abdulaziz, 2003). French-influenced Creoles appeared in Mauritius and Réunion. Regional varieties of former colonial languages are also emerging in urban centers where these languages are appropriated in counter-hegemonic ways (Owino, 2002). This process has reconfigured linguistic ecologies in terms of the relationship between the local and the former colonial languages. The latter are no longer 'foreign' and are more likely *en route* to becoming languages of wider communication in modern metropolitan centers, for example, in Dakar, Maputo, and Johannesburg, they are seldom heard and used on the periphery, however, where they remain quite foreign.

Each of the colonial powers developed language policies that matched their style of administration: direct rule in the case of the French, Spanish, and Portuguese and indirect rule in the case of the Germans and British. Language education policy set in place a pattern that prioritized the use of the colonial languages as the language education target and under-emphasized local languages. For the last century, this policy and practice have been recognized by scholars as detrimental to social, economic, and educational development. Even though most countries achieved their independence during the 1960s, few have made much progress toward reprioritizing local languages in education and in practice. Gramsci (cited in Hoare and Smith, 1971) and Bourdieu (1991) provided the ideological explanations about why it is so difficult for, and not in the interests of, post-colonial governments to alter the existing 'habitus' or status quo.

To add to the complexity, some African languages have also come to displace or diminish the status and functional use of smaller indigenous languages. Arabic (Ayari, 1996), Amharic, and Kiswahili were thus used before, during, and after the colonial

period. Others, like Chichewa (in Malawi) and Setswana (in Botswana) have been used in language policies after independence to achieve the same ends. With few exceptions, though, the emphasis remains directed toward English, French, Portuguese, Spanish, and Arabic, especially in the upper echelons of education, political activity, and the economy. Unfortunately, these languages remain inaccessible to most people in those areas where they now represent the key to power. Formal education, since the late 19th century, has been exclusionary, resulting in large-scale failure (e.g., Obanya, 1999; Bamgbose, 2000; Wolff, 2000; Alidou, 2004). Because language policy apparently serves the interests of the ruling elites, rather than the proletariat, it has also become highly politicized (Mateene, 1999; Ouane, 2003).

Language Policy during the Colonial Period

Direct rule and the prioritization of the colonial languages "facilitated the policy of cultural assimilation, especially in the French, Portuguese and Spanish territories, resulting in the withdrawal from, even denial of, the authentic sources of African creative genius by the new elite" (Boahen, 2003: 238). The French, determined to establish francophone systems, forbade the use of African languages and Arabic in schools: "Literacy through the use of Ajami, Tifinagh, or any other writing system was strictly prohibited" (Alidou, 2004: 199). The 1944 French African Conference in Brazzaville (Diop *et al.,* 2003) reinforced this policy. Written resources in local languages available before colonization were put aside, and further language development in the French colonies was minimal (Abdulaziz, 2003). The concept of education became synonymous with education in French. Similar language policies that discouraged local language use were adopted in the Portuguese and Spanish colonies, although education was less of a priority in these colonies and was left in the hands of Catholic missionaries.

The Germans, Belgians, and British, preferring indirect rule, encouraged the use of indigenous languages to varying degrees. Where African languages were used in education by the British and German administrations, they were the larger, regional lingua francas, rather than the smaller local languages, and were incorporated into transition to German or English programs.

The spread of Kiswahili, which was already a language of trade in East Africa, was advanced during German occupation between 1885 and 1914. German administrators believed it would be useful as a language for governance, and missionaries were encouraged to use Kiswahili for primary education. When Britain took control of what was then Tanganyika in 1918, the use of Kiswahili medium in primary schools was continued, although English was to be used in education from secondary school onward. The British administration thus cemented a late-transition model. At independence in 1961, Tanzania became one of the very few countries to declare an African language, Kiswahili, as its national and official language. German administration left less of an imprint on Namibia where the education of African children was left to the missionaries.

The British, too, preferred indirect rule and encouraged the missionaries' use of African languages in primary schools (about 3 to 4 years, except in Tanzania), with a transition to English in the 4th to 5th year. Some authors attributed Britain's 'weak' support of indigenous languages to the policy of separate development (e.g., Kamanda, 2002), which was most evident in South Africa. Nevertheless, a positive consequence of this policy was the gradual development of orthographies, terminologies, and written texts in many African languages in former British colonies, unlike the situation in territory colonized by Southern European countries. The effects of different policies continue: it takes less time to develop written materials where there are already orthographic systems in place, even when those orthographies are contested.

Language Education Policies and Independence

Even though most countries achieved their independence in the 1960s, the colonial policies continue to the present day. The terms 'Francophone' (including the former Belgian colonies), 'Anglophone,' 'Lusophone' (Portuguese), and 'Arabophone' Africa are frequently found in the literature. They denote the linguistic sphere of influence of the more resilient elements of the 19th–20th century colonial control of Africa, as well as the spread of Arabic that predated and outlasted direct European influence. The first three terms are unsatisfactory, however, because they perpetuate the effect of colonial activity even after independence. Second, given the small percentage of people who speak French, Portuguese, and English proficiently (10–20% according to Heine, 1992; see also Roy-Campbell and Qorro, 1997 and Kamanda, 2002), the terms are misnomers. Third, there is some overlap between Francophone and Arabophone Africa, because Tunisia, Algeria, and part of Morocco were at one time under the control of the French and the French language, but are now administered through Arabic. Fourth, these terms exclude those countries that have developed their own rather different language policies, such as Ethiopia

and Somalia. Both of these countries developed language policies that are strongly oriented toward a single African language, namely Amharic and Somali, respectively.

During the colonial period, a series of educational commissions of inquiry and reports, beginning with the United Missionary Conference in Kenya in 1909 and including the 1953 UNESCO *Report on the use of the vernacular languages in education,* were undertaken. In contrast to existing language policies, each of these reports recommended the use of the mother tongue as the primary medium of education and during the better part of primary school. The recommendations of these early reports were mirrored in a later series of reports and resolutions issued by ministers of education and heads of state, such as the Organization of African Unity Language Plan of Action for Africa (OAU, 1986) and the Asmara Declaration in 2000.

Therefore, upon attaining independence, one might have anticipated a reprioritization of African languages, as well as an attempt to rediscover what worked well in education prior to colonization. Yet, with the exception of Tanzania, Guinea Bissau, South Africa, Ethiopia, Somalia, and most recently, Eritrea, African states have not paid much attention to the maintenance or extension of African languages in the school system. Surprisingly, in most of the former British territories, the colonial practice of using the local language as the medium of instruction for up to 4 years was either whittled away to 3 or fewer years or even done away with altogether, as for example in Zambia, Kenya, Uganda, Zimbabwe, and Botswana. Materials, school texts, teacher training programs in the local languages, and terminology lists relevant at levels above the third year in school fell into disuse and gradually disappeared.

Unsurprisingly, by the late 1980s, most of these countries began to identify crises in English-language levels of achievement in schools: The earlier transition to English was not achieving the desired objectives. Tripathi (1990) suggested that, 25 years after independence, English-only education had reduced the percentage of Zambians proficient in English. Yet, it has taken another 15 years for the reintroduction of mother-tongue education, albeit in a limited and early transition to English form. A protracted time lag between problem identification, policy declaration, and change of practice is fairly common. Malawi's post-independence policy of using one national language – Chichewa (Nyanja) – as the only medium of instruction for the first 4 years of education was revised in 1996 to include other local languages, yet partial implementation began only in 2004.

Little changed in the Spanish, Portuguese, and French colonies during the first decade or more of independence: most formal education continued in French, Portuguese, or Spanish. However, there were several attempts to experiment with various forms of transitional local language (mother-tongue) programs. By 1973, experimental schools were established in Niger and have continued since then to be known as "Bilingual French-National Language Schools" (Wolff, 2000). Since 1975, steps have been taken to transcribe Malian languages and develop orthographies for use in schools in the transition-to-French experimental programs (Hutchison, 1990). Mozambique, which achieved independence in 1975, implemented experimental bilingual school and adult programs (e.g., Benson, 2000; Veloso, 2002) and altered its language education policy in 1996; however, it was not until 2003 that its schools began providing limited use of local languages for 3 years of schooling. A consortium of nongovernmental organizations (NGOs) has collaborated in the Propelca project in Cameroon since 1981 to develop early mother-tongue transition programs and to experiment with a three-language model: the mother tongue, a regional Cameroonian language, and either French or English (Tadajeu and Chiatoh, 2005). The government has not, however, implemented these programs on a widespread basis.

Alidou (2004) noted that, since the 1990 World Conference on Education for All, which was held in Jomtien, there has been increased interest in experimental programs that begin with African languages. Experimental bilingual or trilingual programs that have been usually funded by aid agencies from Germany, the Netherlands, and the Nordic countries have been piloted in several African countries. Well meaning as they may be, these programs are nevertheless always framed as transitional models, with the ultimate target being an exogenous language. Therefore, although these programs attempt to revalue African languages, the process is fundamentally flawed. Models based on the transition from the local language to one of higher status, in which most teachers are not proficient and which the communities seldom hear or use in everyday activities, have already failed across the continent for 120 years (cf. Wolff, 2000; Stroud, 2002). The answer lies in other approaches.

Language policy in the Maghreb took a different course after independence. The spread of Islam from the 7th century onward across North Africa and the establishment of Koranic schools had a significant impact on linguistic practices. In addition to the diglossic situations that developed in relation to Classical (written) Arabic for religious purposes and the numerous informal and spoken varieties of

Moroccan, Tunisian, and Algerian Arabic, for example, there were also numerous varieties of Berber that pre-dated Islam. The Maghreb, including Mauritania and Libya, embarked on an arabization policy after independence. Arabization was intended to gradually dissolve the influence of French, as well as Spanish for part of Morocco and Italian in Libya. The spread of Islam, followed by francophone and then arabization policies, has had a deleterious effect on Berber languages. Although Berber-speaking communities constitute between 45% (Spolsky, 2004) and 60% (Saíb, 2001) of the population in Morocco, and between 20% (Spolsky, 2004: 135) and one-third (Bamgbose, 2000) of the population in Algeria, they have been ignored in policy provision. Even though Modern Standard Arabic has been selected for use as the written language in schools, it is also sufficiently distant from the spoken varieties to create difficulties (Daoud, 2001; Saíb, 2001). Despite political attempts to replace French in former French colonies in the Maghreb, it continues to occupy a prominent position in education, and arabization has not yet achieved dramatic improvements in education. Ayari (1996) ascribed very low levels of literacy in the Maghreb to the doule bind of diglossia and bilingual education that perpetuates the belief and reality that upward mobility is achieved only through an exogenous language.

Language Education Policies with Potential

There have been examples of language policies in Africa that have had successful academic outcomes, and these policies offer useful insights. To make best use of these examples, educators, language specialists, and political groups need to be able to separate educational from political issues.

Somalia is the only country to have achieved, since independence, the use of an African language throughout the years of school education. Government policy and commitment toward rapid implementation of this language policy led to the development of the Somali orthography, the use of Somali in the schools, and the development of educational resources in Somali between 1973 and 1986. What was required was a single-minded government-driven process which included a mass literacy campaign for approximately 2 years during the mid-1970s while other activities were interrupted. Increasing civil strife from 1987 and descent into civil war by 1991, however, has brought about a collapse in the education system, and thus the policy has been short-lived (Griefenow-Mewis, 2004).

Since independence, Ethiopia and Tanzania have each selected and used a dominant African language, Amharic and Kiswahili, respectively, as a medium of instruction throughout the primary school years, followed by the transition to English in secondary school. Although the implementation of these language policies has been accomplished, the weakness of both is that other indigenous languages have been ignored. In each case, few students have found their way into secondary school education and beyond. The response in Tanzania has been pressure to introduce English as a medium of instruction earlier (Roy-Campbell and Qorro, 1997). Ethiopia has, however, chosen to introduce other significant regional languages for the medium of instruction in primary schools beginning in 1994. Implementation has been complicated and slow, and it is too early to assess the efficacy of this alternative (Bloor and Tamrat, 1996). A similar move in Malawi to expand the use of a single national language, Chichewa, by including other local languages, alongside English, indicates a growing awareness that a monolingual African paradigm does not work well in multilingual African societies.

In South Africa and Namibia, there has been local language development and use throughout primary education. The South African government, during the first phase of apartheid (1955–1976), succeeded in training teachers, developing terminology, translating school textbooks, and establishing the mother tongue as the medium of education for 8 years of primary schooling for African pupils in South Africa and Namibia (formerly South West Africa). This was achieved with minimal expenditure: far less was spent on the education of African children than was spent on English- and Afrikaans-speaking learners. Political resistance to apartheid and the compulsory use of Afrikaans alongside English in secondary school resulted in a student uprising in Soweto in June, 1976. Thereafter, mother-tongue education was reduced to 4 years followed by a switch to English for most speakers of African languages. In 1997, a new government, having declared 11 languages as official languages in a democratic constitution, introduced further changes. Although this new language policy facilitated the greater use of African languages, the national department of education ignored it and has established, instead, an even earlier transition to English. The students who have African home languages (some 78%) switch to English after 3 years, whereas ironically, English and Afrikaans speakers continue to use their mother tongue as the medium of education as they did during the colonial and apartheid eras. The gap in educational achievement, predictably, continues to widen (Heugh, 2003).

Between 1966–1984, the first independent government of Guinea Conakry, formerly a French colony, also established mother-tongue education in eight (later reduced to six) local languages for 8 years of

schooling. This policy seems to have failed because parents and communities were not sufficiently consulted, and the process was entirely top-down, linked specifically to the autocratic leadership of Sékou Touré (Kazadi, 1991; Yerende, 2005). Political changes since 1984 resulted in a reversal of the policy to the use of French as the medium of instruction. There has been recent discussion of reintroducing limited mother-tongue education by adapting a model from Mali in which 3 years of schooling in the mother tongue is followed by a transition to French (Yerende, 2005).

The Six Year Primary Project in Ife, Nigeria, in which Yoruba was used as the medium of instruction for the 6 years of primary education, has been well documented (e.g., Bamgbose, 2000; Adegbija, 2003). The significant lesson from this longitudinal study is that use of the mother tongue as the medium of instruction does not prevent the achievement of proficiency in the language of wider communication (LWC). Rather, the extended use of the mother tongue facilitates high levels of proficiency in the other target language. This project confirms the current theories of additive bilingual education in an African setting.

The Somali, South African, and Guinean examples reveal specific issues that could inform future developments. First, where there is political will, provision of education in the mother tongue or local language development is feasible in Africa and can be implemented within a short period of time. Second, each case shows that cost is not a prohibitive factor. These achievements occurred with minimal expenditure. The third lesson, from each of these 3 countries is that if the communities have not been adequately consulted and policy implementation is linked to an undemocratic government, policy success is short-lived.

The Nigerian and South African cases provide other critical evidence. African children achieved considerable academic success during both the Six Year Yoruba program (Bamgbose, 2000) and the 8 years mother-tongue program followed by transition to English and sometimes Afrikaans (Heugh, 2003). In each case, the mother tongue was used as medium of instruction, and English was taught as a subject only, but by teachers proficient in the language. By the time students were required to switch to English, they had kept up with the academic requirements of the rest of the curriculum, and they had sufficient proficiency to switch languages. The common denominator in these examples is 6 or more years of use of the mother tongue as the medium of instruction before making a transition/switch to an exogenous language. The only examples of successful early transition models in Africa have been in the late 19th- and 20th-century mission schools and more recent private school systems that cater to the children of the elites. Submersion and early transition models that are predominantly found in Africa, which affect the majority of those who are able to attend school, fail.

Residual Effects of Colonial Policies

The experience of colonization and the education of a small African elite that was to be assimilated into the administrative apparatus effectively undermined confidence in African languages as adequate conduits of education. Pre-colonial education systems in Ethiopia and the Sahel have shown that African languages can be used as languages of education through tertiary education. Post-colonial attempts to make systemic educational use of African languages in South Africa, Somalia and Guinea-Conakry have been accompanied by repressive government and thus discontinued after political change.

For the most part, post-colonial education systems have continued colonial policies and generally underutilized African languages and knowledge systems. They have not produced satisfactory academic achievement and are characterized by very high dropout and repeater rates. A UNESCO study conducted in 2000 of literacy and numeracy levels of fourth-grade students in African countries showed consistently poor results across Africa (an average of 32% for numeracy and 25% for literacy in Nigeria; Alidou, 2004). The majority of children who enter school do not go beyond primary schooling. Colonial education and the use of colonial languages have increased social inequalities along class and gender lines (Obanya, 1999; Bamgbose, 2000; Ouane, 2003; Alidou, 2004).

Why Are the Colonial Languages Retained in Education?

One reason that the former colonial language is retained for high status functions is that it is perceived as neutral. This perception has been advanced notably in Zambia, Namibia, South Africa, and Mozambique. The existence of multiple local languages is often thought to increase the likelihood of ethnic rivalry and tensions and civil strife. However, recent examples of ethnic conflict in Somalia, Rwanda, and Burundi show that it occurs even when people speak the same language. Another argument is that the management of multilingualism is believed to be more expensive than that of monolingualism in the colonial language. However, this argument does not stand up to empirical data (Grin, 2005); the costs of not providing mother-tongue education (school failure, high dropout rates, and alienated youth)

are likely to be higher. Most authors have argued that keeping the colonial language for high-status functions without simultaneously promoting African languages is to the advantage only of the ruling elite.

Conclusion

Every year, at one conference after another, language experts identify the ill-fitting and misapplied language policies of the African continent, and they resolve to persuade governments to change course. Changes do occur, but they are too little and too slow. Programs that include more African languages will not change the current educational prognosis if they are designed merely as a transition to Arabic, Spanish, French, Portuguese, or English. Another solution needs to be found, one that requires a completely new approach to language acquisition theory and practice in Africa. The answer may lie in the language chains and informal acquisition of multilingualism that lubricate communication across the continent, rather than with governments' capacity to change and effect appropriate policy.

See also: Africa as a Linguistic Area; African Linguistics: History; Amharic; Applied Linguistics in Africa; Arabic; Bamgbose, Ayo (b. 1932); Berber; Bilingual Education; Bilingualism and Second Language Learning; Bilingualism; Cape Verdean Creole; Education in a Former Colonial Language; Education in a Multilingual Society; Egypt: Scripts, Ancient; English in the Present Day (since ca. 1900); Ethiopian Semitic Languages; French; Ge'ez; Hausa; Language Planning and Policy: Models; Language Policy in Multilingual Educational Contexts; Lingua Francas as Second Languages; Missionary Linguistics; Mother Tongue Education: Nonstandard Language; Mother Tongue Education: Standard Language; Pidgins and Creoles: Overview; Portuguese; Somali; Spanish; Swahili; Tadadjeu, Maurice (b. 1950); Tigrinya; Wolof; Yoruba.

Bibliography

Abdulaziz M H (2003). 'The history of language policy in Africa with reference to language choice in education.' In Ouane (ed.). 181–199.

Adegbija E (2003). 'Central language issues in literacy and basic education: three mother tongue education experiments in Nigeria.' In Ouane (ed.). 299–331.

Alidou H (2004). 'Medium of instruction in post-colonial Africa.' In Tollefson J W & Tsui A B (eds.) *Medium of instruction policies: which agenda? whose agenda?* Mahwah, NJ & London: Lawrence Erlbaum. 195–214.

Ayari S (1996). 'Diglossia and illiteracy in the Arab world.' *Language, Culture and Curriculum 9(3),* 243–253.

Bamgbose A (1994). 'Pride and predjudice in multilingualism.' In Fardon R & Furniss G (eds.) *African languages, development and the state.* London & New York: Routledge. 33–43.

Bamgbose A (2000). *Language and exclusion: the consequences of language policies in Africa.* Münster, Hamburg & London: Lit Verlag.

Battestini S (1997). *Ecriture et Texte: Contribution Africaine. Québec & Ottawa: Les Presses de l'Université Laval.* Paris: Présence Africaine.

Benson C (2000). 'The primary bilingual education experiment in Mozambique, 1993–1997.' *International Journal of Bilingual Education and Bilingualism 3(3),* 149–166.

Bloor T & Tamrat W (1996). 'Issues in Ethiopian language policy and education.' *Journal of Multilingual and Multicultural Development 17(5),* 321–338.

Boahen A (ed.) (2003). *General history of Africa. VII: Africa under colonial domination 1880–1935* (abridged edn.). Cape Town: NAE & UNESCO.

Bourdieu P (1991). *Language and symbolic power.* Cambridge: Polity Press.

Chumbow B S (1999). 'Transborder languages of Africa.' *Social Dynamics 25(1),* 51–69.

Daoud M (2001). 'The language situation in Tunisia.' *Current Issues in Language Planning 2(1),* 1–52.

Diop M, Birmingham D, Hrbek I, Margarido A & Niane D (2003). 'Tropical and equatorial Africa under French, Portuguese and Spanish domination, 1935–45.' In Mazrui A & Wondji C (eds.) *General history of Africa. VIII: Africa since 1935* (unabridged edn.). Cape Town: NAE & UNESCO. 58–75.

Djian J-M (2004). 'Mali: the fabulous past of Africa.' *Le Monde diplomatique,* August, 10–11.

Djité P G (1993). 'Language and development in Africa.' *International Journal of the Sociology of Language 100/101,* 149–166.

Elaigwu J & Mazrui A (2003). 'Nation-building and changing political structures.' In Mazrui A & Wondji C (eds.) *General history of Africa. VIII: Africa since 1935* (unabridged edn.). Cape Town: NAE & UNESCO. 435–467.

Fardon R & Furniss G (1994). 'Introduction: frontiers and boundaries – African languages as political environment.' In Fardon R & Furniss G (eds.) *African languages, development and the state.* London & New York: Routledge. 1–29.

Griefenow-Mewis C (2004). 'How to turn oral African languages into written languages? The Somali and Oromo experience.' In Pfaffe J F (ed.) *Making multilingual education a reality for all: operationalising good intentions. Proceedings of the Joint Third International Conference of the Association for the Development of African Languages in Education, Science and Technology (ADALEST) and the Fifth Malawian National Language Symposium.* Zomba: Centre for Language Studies, University of Malawi & GTZ. 204–211.

Grin F (2005). 'The economics of language policy implementation: identifying and measuring costs.' In Alexander N (ed.) *Mother Tongue-Based Education in Southern Africa:*

The Dynamics of Implementation. Cape Town: PRAESA. 11–25.

Heine B (1992). 'Language policies in Africa.' In Herbert R K (ed.) *Language and society in Africa*. Johannesburg: Witwatersrand University Press. 23–35.

Heugh K (2003). 'A re-take on bilingual education in and for South Africa.' In Fraurud K & Hyltenstam K (eds.) *Multilingualism in Global and Local Perspectives. Selected papers from the 8th Nordic Conference on Bilingualism*. Stockholm: Centre for Research on Bilingualism, Stockholm University and Rinkeby Institute of Multilingual Research. 47–62.

Hoare Q & Smith G N (eds. & transl.) (1971). *Antonio Gramsci: selections from the prison notebooks*. London: Lawrence & Wishart.

Hutchison J (1990). *Final report: evaluation of the Experimentation in National Languages in Primary Education in the Republic of Mali*. Washington, DC: USAID.

Kamanda M C (2002). 'Mother tongue education and transitional literacy in Sierra Leone: prospects and challenges in the 21st century.' *Language and Education: An International Journal 16(2)*, 195–211.

Kazadi N (1991). Le français et/ou les langues nationales. In *L'Afrique afro-francophone. Collection Langues et développement*. Paris: Didier Érudition. 135–137.

Ki-Zerbo J (ed.) (2003). *General history of Africa. 1: Methodology and African prehistory* (abridged edn.). Cape Town: NEA & UNESCO.

Makoni S (1998). 'African languages as European scripts: the shaping of communal memory.' In Nuttall S & Coetzee C (eds.) *Negotiating the past: the making of memory in South Africa*. Cape Town: Oxford University Press. 242–248.

Mateene K (1999). 'OAU's strategy for linguistic unity and multilingual education.' *Social Dynamics 25(1)*, 164–178.

OAU (1986). *Language plan of action for Africa*. Addis Ababa: Organisation of African Unity.

Obanya P (1999). *The dilemma of education in Africa*. Dakar: UNESCO Regional Office.

Ouane A (2003). 'Introduction: the view from inside the linguistic jail.' In Ouane A (ed.) *Towards a multilingual culture of education*. Hamburg: UNESCO Institute for Education. 1–22.

Owino F (2002). 'Conquering the conqueror. The empowerment of the African languages casts a shadow over English in Africa.' *Perspectives in Education 20(1)*, 197–212.

Ranger T (1989). 'Missionaries, migrants and Manyika. The invention of ethnicity in Zimbabwe.' In Vail L (ed.) *The creation of tribalism in Southern Africa*. London/Berkeley & Los Angeles: James Currey/University of California Press. 118–150.

Roy-Campbell Z M & Qorro M A (1997). *Language crisis in Tanzania: the myth of English versus education*. Dar es Salaam: Mkuki Na Nyota Publishers.

Saíb J (2001). 'Berber and Arabic in Morocco.' In Extra G & Gorter D (eds.) *The other languages in Europe*. Clevedon, Buffalo, Toronto & Sydney: Multilingual Matters. 429–444.

Sow A & Abdulaziz M (2003). 'Language and social change.' In Mazrui A & Wondji C (eds.) *General history of Africa. VIII: Africa since 1935*. Cape Town: NAE & UNESCO. 522–552.

Spolsky B (2004). *Language policy*. Cambridge: CUP.

Stroud C (2002). *New Education Division documents 10: towards a policy of bilingual education in developing countries*. Stockholm: SIDA.

Tadajeu M & Chiatoh B (2005). 'Mother tongue-focused bilingual education in Cameroon.' In Alexander N (ed.) *Mother Tongue-Based Education in Southern Africa: The Dynamics of Implementation*. Cape Town: PRAESA. 123–136.

Tripathi P D (1990). 'English in Zambia: the nature and prospects of one of Africa's 'new Englishes.' *English Today 6(3)*, 34–38.

Veloso T (2002). 'Becoming literate in Mozambique – the early stages in Sena (Cisena) and Shangaan (Xichangana).' *Perspectives in Education 20(1)*, 97–109.

Wolff E H (2000). 'Indigenous languages in education: learning from West African experiences.' In Pfaffe J F (ed.) *Local languages in education, science and technology. Proceedings of the Second National Symposium on Language Policy Formulation*. Zomba: University of Malawi & GTZ. 52–71.

Yerende I E (2005). 'Ideologies of language and schooling in Guinea-Conakry: a postcolonial experience.' In Brock-Utne B & Hopson R K (eds.) *Languages of instruction for African emancipation: focus on postcolonial contexts and considerations*. CASAS Book Series No. 39. Cape Town: CASAS. 119–230.

Language Education Policies in North America

D Christian, Center for Applied Linguistics, Washington, DC, USA

Naming the official language(s) is the most readily recognized manifestation of language policy. In North America, Canada and the United States differ markedly in this respect. The United States has no explicitly declared official language (although throughout its history, there have been attempts to amend the United States Constitution to make English the official language). In contrast, Canada has an official policy of bilingualism, with the Official Languages Act (1969) declaring French and English to be the nation's official languages. The scope of language policy extends beyond designating the official languages, however, to "the combination of official decisions and prevailing public practices related to language education and use" (McGroarty, 1997: 67). In fact, Spolsky (2004: 46) maintains, "Of all the domains for language policy, one of the most important is the school." Thus, language education policies can be viewed as a significant subset of language policy broadly defined.

This article examines the language education policies of the United States and Canada, with the official language policy of the two nations (or lack of it) as backdrop. In both countries, education is primarily within the jurisdiction of the states or provinces, rather than the federal government, but policies are, in fact, framed at multiple levels – federal, state or provincial, and local – and they are enacted through executive, legislative, and judicial actions at each of these levels. Given the large number of state and local jurisdictions represented in the two countries, this article will focus on language education policies at the federal level, with occasional reference to significant state policies.

Language Education Policies

Both Canada and the United States have diverse populations, with many languages spoken in communities around each country. In the United States, students from non-English-speaking backgrounds represent the fastest-growing segment of the student population and constituted over 10% of the K–12 public school student population in 2003 (Padolsky, 2004). These students speak over 400 languages, but nearly 80% are native Spanish speakers (Kindler, 2002). In Canada, in addition to the two official languages, communities use "11 different Aboriginal language families... and over 100 other languages actively used by immigrants and their descendants" (Burnaby, 2002). As a result, language education policies (implicit or explicit) cover at least three areas of concern:

- Medium of instruction issues
- Heritage languages development
- Learning of new languages for all students.

Policies in the United States

Broadly speaking, education policies in the United States express the following views on language:

- Although there is no official language, proficiency in (standard) English is widely recognized as necessary, desirable, and a goal of education.
- Learning languages other than English is encouraged for instrumental reasons related to national defense and economic competitiveness, with language education supported in limited ways.
- Services in languages other than English for nonnative speakers are encouraged in some educational contexts, to allow access to the educational system, but maintenance of heritage languages is not generally supported (Christian, 1999).

In general, U.S. language education policies treat *language diversity as a problem* to be addressed, whether it is diversity among residents of this country or diversity around the world among those we may need to communicate with (as opposed to viewing *languages as resources*; see Ruiz (1988) for a discussion of this contrast).

The Native American Languages Act passed in 1990 is an exception to the preceding in that it declares that U.S. policy is to "preserve, protect and promote the rights and freedom of Native Americans to use, practice, and develop Native American languages." Although this act does not represent language education policy directly, it can be used to justify learning Native American languages in schools (McCarty, 2004).

Speakers of Languages Other Than English

Language education policies related to medium of instruction and maintenance of native languages other than English has become highly salient with increasing linguistic diversity in the schools. Discussion of these policies is often contentious because of mixed attitudes toward pluralism and their overlap with many highly charged political issues such as immigration. As a result, proposals for legislation for K–12 as well as adult education that relate to this population have been hotly debated.

Federal law has never required that native languages be used as medium of instruction. Education laws since 1968 have authorized funding to help schools develop programs to assist students with limited English proficiency to learn English and succeed academically, primarily through programs that facilitate transition to an all-English curriculum. A Supreme Court decision (*Lau v. Nichols* in 1974) held that schools violate students' civil rights when they fail to provide students from a non-English-language background with special services that help them to participate more fully in public education (but medium of instruction was not specifically addressed). Since that time, the major mechanisms for language education policy have been through the courts on civil rights bases and through appropriations determining the amount and targeting of funding for states, school districts, and schools. The most recent major education law, the No Child Left Behind Act of 2002 does not enact language education policy. It explicitly makes schools responsible for the academic progress of all groups of students and requires that English language proficiency outcomes demonstrate annual progress for nonnative speakers. Although native language development is neither prohibited nor encouraged, the emphasis on performance in English and on accountability for schools and students places a premium on English-medium instruction in current policy and practice. Thus, policies on the use of languages other than English in instruction (to serve students from non-English-speaking backgrounds) have not been intentionally or specifically formulated at the federal level, but have evolved from a series of legislative, judicial, and regulatory actions.

States vary widely in their policies on medium of instruction. Since the 1960s, some have required use of native language in schools where a certain number of students speak the same language. However, several widely publicized campaigns since the mid-1990s have culminated in public votes on bilingual education (Crawford, 2004). In three cases (California, Arizona, and Massachusetts), the referendum passed and the practice of bilingual education was severely constrained. In one case (Colorado), the measure did not pass.

Learning Additional Languages

Another major domain of language education policy covers the learning of languages other than English (often termed "foreign" languages, "second" languages, or "world" languages). This domain, for example, includes the designation of high-priority languages for the country and taking actions to make sure that the needed programs for learning additional languages exist. In the United States, there has been a general policy direction of broadening the number of languages available for study, particularly among "less commonly taught" languages, through funding for university-based centers and fellowships, materials and curriculum development, and study abroad programs. Lists of "critical" languages are periodically generated for different purposes by the Departments of Education or Defense, and legislative and regulatory proposals specify particular languages for attention.

Policy support in the United States for foreign languages has been somewhat stronger in higher education than at the elementary and secondary levels, but it remains far from a mandate or requirement. Less than 10% of college students study a language, and less than 10% of those students learn a less commonly taught language (NSEP, 2003). In the 1950s, the National Defense Education Act included the goal of strengthening modern foreign language instruction, with priority given to languages judged to be needed by government, business, and industry. Later, the provisions moved to the Higher Education Act, with a focus on building capacity in the system through such mechanisms as funding for university resource centers, fellowships, undergraduate international studies and foreign language programs, business and international education programs, and international research and studies program.

Another program reflects more recent directions in federal language education policy. The National Security Education Program, established in 1991 and administered through the Department of Defense, provides funding for study-abroad opportunities for undergraduate and graduate students, with special focus on nonlanguage specialists, and for projects to improve international and language education in institutions of higher education (NSEP, 2003). NSEP emphasizes non-Western European languages and exemplifies the federal priority on expanding the range of languages being studied.

In June 2004, The National Language Conference was convened by the U.S. Department of Defense to "to identify actions that could move the United States toward a language-competent nation" (www.nlconference.org). Participants from government, military, education, and private sectors assessed the country's needs and issued a call to action to improve its language capacity. The conference took a clearly instrumental orientation toward valuing skills in languages other than English (for defense and economic competitiveness). Future policy actions related to learning languages other than English may flow from the results of this activity expressed in a white paper "recommending the initial steps toward a national language agenda" that is in preparation.

In summary, policy toward the learning of foreign languages takes a very pragmatic, instrumental approach. As concerns over national security have grown, the role of language education to address those issues has been the driving force of policy.

Policies in Canada

In Canada, the context of official bilingualism changes the policy landscape somewhat, in that both French and English are the major societal (and official) languages. As in the United States, education in Canada is governed by the provinces, so the federal government has no national level education (or language education) policy, except as related to the official languages. French and English are protected as languages of instruction for their native speakers, and they are the two languages that are taught as second languages in most schools in Canada.

Medium of Instruction

According to Cumming (2000), speakers of English or French in a context where the other official language dominates have "official minority" status (e.g., English in Quebec or French in Ontario). Federal policy guarantees children with an official minority language background (French or English) the right to elementary and secondary education in their native language. Immigrant and indigenous families who speak languages other than French or English do not have similar rights, although in some provinces, use of local community languages as the medium of instruction is allowed (Burnaby, 2002).

Speakers of Languages Other Than French and English

Language education policies for speakers of languages other than French or English are primarily set at the provincial level. As in the United States, the focus is on learning one of the official languages, with native language as medium of instruction permitted but not required. Some attention has been given to maintaining heritage languages in Canada, particularly for the indigenous (Native or First Nations) populations. Some provinces have developed programs to promote specific Native languages, including use of the languages as medium of instruction (e.g., Cree; see Burnaby, 2002). For the languages of immigrants, such as Chinese or Ukrainian, about half the provinces have curricula that include instruction in the local heritage languages, but it is primarily outside the official school day. Some exceptions, such as Ukrainian immersion programs in Manitoba and Alberta, do exist (Cumming, 2000).

Learning Additional Languages

As a result of the official language policy, French (in English-dominant communities) and English (in French-dominant communities) are the most commonly taught additional languages in Canada, from elementary through higher education. Several bilingual French–English universities exist around the country. Government estimates put the percentage of bilingual Canadians at 18%, and a plan was launched recently to increase that number, aiming to provide "new momentum for Canada's linguistic duality" (Government of Canada, 2003). This five-year Action Plan for education and other public sectors promotes the use of both official languages and sets a goal that, by 2013, half of all secondary school graduates will be bilingual (English and French). The plan is backed by increased funding to help schools and communities achieve these goals.

Language education policies with respect to other languages vary from province to province, including the age at which language study begins. Many languages are offered as school subjects, including classical and modern languages, as well as languages of immigrant and indigenous populations.

Conclusion

In summary, neither the United States nor Canada has federal-level, explicit language education policies (with the exception of attention to French and English as official languages in Canada), but a variety of legislative, funding, and regulatory actions constitute implicit policies. In underlying orientation, however, there are some differences. Despite their diffuseness, implicit U.S. language education policies share an orientation to *language as problem* and learning languages primarily for instrumental reasons. In contrast, Canada pursues some language education policies that value language as a resource, particularly in the case of French and English (evidenced by the recent strategy to increase the incidence of bilingualism) and to a lesser extent by a commitment to other heritage languages, both indigenous and immigrant.

See also: Bilingual Education; Bilingualism and Second Language Learning; Bilingualism; Canada: Language Situation; Education in a Multilingual Society; Foreign Language Teaching Policy; Language Policy in Multilingual Educational Contexts; Minorities and Language; Minority Languages: Education; Mother Tongue Education: Standard Language; Multiculturalism and Language; United States of America: Language Situation.

Bibliography

Burnaby B (2002). 'Reflections on language policies in Canada: three examples.' In Tollefson J W (ed.) *Language policies in education: critical issues*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers. 65–86.

Christian D (1999). 'Looking at federal education legislation from a language policy/planning perspective.' In Huebner T & Davis K (eds.) *Sociopolitical perspectives on language policy and planning in the USA*. Amsterdam: John Benjamins Publishing Company. 117–130.

Crawford J (2004). *Educating English learners: language diversity in the classroom*. Los Angeles, CA: Bilingual Educational Services, Inc.

Cumming A (2000). *Second language education in schools in Canada*. Technical report. Available at: http://www.oise.utoronto.ca/MLC/pufahlrep.pdf.

Government of Canada (2003). *The next act: new momentum for Canada's linguistic duality. The action plan for official languages*. Ottawa: Privy Council Office, Government of Canada. Available at: http://www.pco-bcp.gc.ca/aia/default.asp?Language = E&Page = ActionPlan.

Kindler A L (2002). *Survey of the states' limited English proficient students and available educational programs and services: 2000–2001 summary report*. Washington, DC: National Clearinghouse for English Language Acquisition.

McCarty T (2004). 'Dangerous difference: a critical-historical analysis of language education policies in the United States.' In Tollefson J W & Tsui A B M (eds.) *Medium of instruction policies: which agenda? whose agenda?* Mahwah, NJ: Lawrence Erlbaum Associates. 71–93.

McGroarty M (1997). 'Language policy in the USA: national values, local loyalties, pragmatic pressures.' In Eggington W & Wren H (eds.) *Language policy: dominant English, pluralist challenges*. Amsterdam: John Benjamins Publishing Company. 67–90.

National Security Education Program (2003). Report to Congress, September 2003. Available at: http://www.ndu.edu/nsep/NSEPRpt.pdf.

Padolsky D (2004). *How many school-aged English language learners (ELLs) are there in the U.S.?* National Clearinghouse for English Language Acquisition. Retrieved November 11, 2004 from http://www.ncela.gwu.edu/expert/faq/01leps.htm.

Ruiz R (1988). 'Orientations in language planning.' In McKay S & Wong S L (eds.) *Language diversity: problem or resource?* New York: Newbury House. 3–25.

Spolsky B (2004). *Language policy*. Cambridge: Cambridge University Press.

Language Education Policies in South Asia

Y Kachru, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Languages of administration and education have always been matters of concern to the state in all parts of the world. Conquests, empires, and religious proselytization have been powerful agents of the spread of classical languages such as Arabic, Chinese, Greek, Latin, Pali, Persian (Farsi), and Sanskrit and of modern languages such as English, Spanish, and Portuguese. The states of South Asia have been engaged in determining and revising their language and language-education policies since obtaining their independence from colonial rule in the mid 20th century. This is a brief sketch of the language education policies of India, Pakistan, Bangladesh, Nepal, and Sri Lanka.

Multilingualism

With more than 330 languages and 'mother tongues' in India (1991 census), 69 languages and dialects in Pakistan (Rahman, 2004a), 37 languages and dialects in Bangladesh, more than 18 languages in Nepal (Savada, 1993), and 5 languages and creoles in Sri Lanka, South Asia is home to a great many of the world's languages. Virtually all of the people of the region are at least bilingual, and many are multilingual. Even speakers of major languages, e.g., Hindi speakers in India, are fluent in a local language or dialect (e.g., Bhojpuri, Magahi, Garhwali) besides Standard Hindi. The same is true of Pakistan, where people may speak Punjabi (Panjabi) or Sindhi in addition to Urdu, and of Nepal, where although Nepali is used widely, in some regions Hindi is used as a second language.

In view of the number of languages spoken in South Asia, the entire region faces the problems of (a) determining the domains of use of each language, (b) allocating resources to all the languages to ensure their relevance to most domains of activity, and (c) forging a national identity not fractured by narrow language loyalties. The attitudes toward and perceptions of English accompanying its spread, functions, and increasing use in national and international domains throughout the area add to this already complex picture.

Language Policy

South Asia is home to four families of languages: Indo-Aryan, Dravidian, Sino-Tibetan, and

Bibliography

Burnaby B (2002). 'Reflections on language policies in Canada: three examples.' In Tollefson J W (ed.) *Language policies in education: critical issues.* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers. 65–86.

Christian D (1999). 'Looking at federal education legislation from a language policy/planning perspective.' In Huebner T & Davis K (eds.) *Sociopolitical perspectives on language policy and planning in the USA.* Amsterdam: John Benjamins Publishing Company. 117–130.

Crawford J (2004). *Educating English learners: language diversity in the classroom.* Los Angeles, CA: Bilingual Educational Services, Inc.

Cumming A (2000). *Second language education in schools in Canada.* Technical report. Available at: http://www.oise.utoronto.ca/MLC/pufahlrep.pdf.

Government of Canada (2003). *The next act: new momentum for Canada's linguistic duality. The action plan for official languages.* Ottawa: Privy Council Office, Government of Canada. Available at: http://www.pco-bcp.gc.ca/aia/default.asp?Language = E&Page = ActionPlan.

Kindler A L (2002). *Survey of the states' limited English proficient students and available educational programs and services: 2000–2001 summary report.* Washington, DC: National Clearinghouse for English Language Acquisition.

McCarty T (2004). 'Dangerous difference: a critical-historical analysis of language education policies in the United States.' In Tollefson J W & Tsui A B M (eds.) *Medium of instruction policies: which agenda? whose agenda?* Mahwah, NJ: Lawrence Erlbaum Associates. 71–93.

McGroarty M (1997). 'Language policy in the USA: national values, local loyalties, pragmatic pressures.' In Eggington W & Wren H (eds.) *Language policy: dominant English, pluralist challenges.* Amsterdam: John Benjamins Publishing Company. 67–90.

National Security Education Program (2003). Report to Congress, September 2003. Available at: http://www.ndu.edu/nsep/NSEPRpt.pdf.

Padolsky D (2004). *How many school-aged English language learners (ELLs) are there in the U.S.?* National Clearinghouse for English Language Acquisition. Retrieved November 11, 2004 from http://www.ncela.gwu.edu/expert/faq/01leps.htm.

Ruiz R (1988). 'Orientations in language planning.' In McKay S & Wong S L (eds.) *Language diversity: problem or resource?* New York: Newbury House. 3–25.

Spolsky B (2004). *Language policy.* Cambridge: Cambridge University Press.

Language Education Policies in South Asia

Y Kachru, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Languages of administration and education have always been matters of concern to the state in all parts of the world. Conquests, empires, and religious proselytization have been powerful agents of the spread of classical languages such as Arabic, Chinese, Greek, Latin, Pali, Persian (Farsi), and Sanskrit and of modern languages such as English, Spanish, and Portuguese. The states of South Asia have been engaged in determining and revising their language and language-education policies since obtaining their independence from colonial rule in the mid 20th century. This is a brief sketch of the language education policies of India, Pakistan, Bangladesh, Nepal, and Sri Lanka.

Multilingualism

With more than 330 languages and 'mother tongues' in India (1991 census), 69 languages and dialects in Pakistan (Rahman, 2004a), 37 languages and dialects in Bangladesh, more than 18 languages in Nepal (Savada, 1993), and 5 languages and creoles in Sri Lanka, South Asia is home to a great many of the world's languages. Virtually all of the people of the region are at least bilingual, and many are multilingual. Even speakers of major languages, e.g., Hindi speakers in India, are fluent in a local language or dialect (e.g., Bhojpuri, Magahi, Garhwali) besides Standard Hindi. The same is true of Pakistan, where people may speak Punjabi (Panjabi) or Sindhi in addition to Urdu, and of Nepal, where although Nepali is used widely, in some regions Hindi is used as a second language.

In view of the number of languages spoken in South Asia, the entire region faces the problems of (a) determining the domains of use of each language, (b) allocating resources to all the languages to ensure their relevance to most domains of activity, and (c) forging a national identity not fractured by narrow language loyalties. The attitudes toward and perceptions of English accompanying its spread, functions, and increasing use in national and international domains throughout the area add to this already complex picture.

Language Policy

South Asia is home to four families of languages: Indo-Aryan, Dravidian, Sino-Tibetan, and

Austro-Asiatic. Speakers of all these families of languages inhabit all the countries in the region except Sri Lanka. India, the largest nation of South Asia, has always been a multilingual state, exhibiting great linguistic and ethnic diversity. The constitution of India (1950) declares India a "Sovereign Socialist Secular Democratic Republic" (Preamble) and recognizes Hindi in the Devanagari Script as the Official Language of the Indian Union and English as the Associate Official Language. Its Eighth Schedule also lists a number of national languages (originally 14, later 15, and by now 22); the Eighth Schedule keeps expanding because inclusion in it bestows certain advantages on the speakers of a named language (Krishnamurti, 1995: 15).

The proposals in the *National curriculum framework for school education: 2000,* prepared by the National Council of Educational Research and Training, recommends that a mother tongue or regional language should be the first language to be studied in the first two years of schooling. In the next three years, one modern Indian language other than the mother tongue or regional language (Hindi in non-Hindi-speaking states) and English should be added. This three-language formula (a regional language, English, and Hindi in non-Hindi-speaking states; Hindi, another modern Indian language, and English in Hindi-speaking states) should continue in the first two years of secondary school. Sanskrit should be introduced as part of the curriculum in regional language or Hindi at the primary level as it has had impact on all languages of India. At the upper secondary stage, an option to study Sanskrit parallel to English or Hindi should be made available. The final decision about introducing English at the primary or early secondary level has been left to the states; all states by now have made English mandatory as a second or third language in the primary or early secondary grades.

The Islamic Republic of Pakistan has gone through several cycles of framing and promulgating a constitution. According to the constitution adopted in 1973 and modified subsequently, Urdu is the official language of Pakistan; English, however, performs the real functions of an official language. The number of languages listed for Pakistan is 69; only a handful of these have any numerical or political strength. Urdu is spoken by only 7% of the population as a first language and 70% of the population as a second language. The numerically strongest group, Punjabi speakers (44.15%), identify with Urdu, unlike Sindhi speakers, who take a great deal of pride in their language and its great literary tradition. Pashto speakers (15.42%) and Sindhi speakers (14.10%) treat their languages as markers of their respective identities (Rahman, 2004a). Both groups, however, have reconciled themselves to education in Urdu because they realize its value in increased mobility and career advancement. As Rahman observes, "the consequence of privileging Urdu strengthens ethnicity while, at the same time and paradoxically, it threatens linguistic and cultural diversity in the country" (2004a: 5).

Language played a significant role in the movement to make Bangladesh a nation separate from Pakistan. Although Bengali, or Bangla as it is called by the speakers of the language, had the largest number of speakers, Urdu was imposed as the sole official language in both West and East Pakistan. This act added to the resentment the Bengalis felt on other grounds, e.g., administrative, economic, and political. Bangla became a symbol and a cause for which the Bengali people were willing to fight, and they achieved their goal of an independent nation in 1971.

The constitution of Bangladesh (adopted in 1972; current version as amended in 1996) declares Bangla the national language of the nation and does not mention any other language. The speakers of Austro-Asiatic, Dravidian, and Sino-Tibetan languages constitute small minorities. The constitution recognizes neither their ethnic identity nor their languages; the emphasis is on strengthening a Bengali culture and identity.

The 1978 constitution of Sri Lanka designates Sinhala the official language, Tamil "also an official language," and English a link language. The constitution states that all official documents and records are to be kept in Sinhala except in the Northern and Eastern provinces, where Tamil is to be used, and that Tamil or English translations of all official document and records are to be made available upon demand. There is provision for education – including university education – in Sinhala, Tamil, or English. The emphasis in chapter 1 of the constitution is on forging 'a Sri Lankan identity, recognizing the multiethnic, multilingual and multireligious character of Sri Lankan society.' Sinhalese, as the official language, and Buddhism, as the religion, are given 'the foremost place' by the constitution.

The national and official language of Nepal, as proclaimed by the Ministry of Education and Sports, His Majesty's Government of Nepal, is Nepali in the Devanagari script. All other languages spoken in Nepal are recognized as languages of the nation. Nepal is home to two distinct groups of languages: Indo-Aryan and Tibeto-Burman. The Indo-Aryan languages with the largest number of speakers are Nepali (12 million), Maithili (2.85 million), and Bhojpuri (1.79 million); the Tibeto-Burman languages with the most speakers are Tamang (1.2 million) and

Newari (1 million). All linguistic groups acquire some competence in the official language, Nepali, through the school system. Hindi or Hindustani functions as a lingua franca in local markets. In spite of the country's linguistic, ethnic, and religious diversity, Nepalis have forged a distinct national identity.

Mediums of Instruction

In all the countries of South Asia, official policy notwithstanding, there is a great deal of variation among different regions in the type and quality of education. In India, following the provision in the constitution regarding linguistic rights, efforts have been made to provide primary education through the mother tongue (Sharma, 1985). Parents, however, prefer the regional or an official language as the medium of instruction for their children in view of the clear advantages these languages provide in terms of mobility and career choices (for a discussion of issues in language, education, and society, see Krishnamurti, 1998; Pattanayak, 1991). Accordingly, the number of languages used as mediums of instruction has varied considerably since the 1950s, and the national and official languages are flourishing as mediums of instruction. Schools that have national languages and Hindi as mediums are generally state supported; most English-medium schools are privately run, many of them by Christian missions (the convent schools), although they receive financial support from the state as long as they follow the national curriculum. Economic factors, overwhelming linguistic diversity, and Christian missionary work have led nine states and Union Territories of India to declare English their official state language (for a list, see Krishnamurti, 1995: 17).

The official language policy in Pakistan has been to encourage the learning of Arabic as the language of religion and Urdu as the medium of instruction for the entire nation in order 'to transform the Pakistani Nation into an integrated, cohesive entity.' The objective is to 'make the Qur'anic principles and Islamic practices as an integral part of curricula so that the message of the Holy Qur'an could be disseminated in the process of education as well as training' (Pakistan.gov). A majority of public schools use Urdu as the medium of instruction; Sindhi and Pashto are confined to their respective regions. However, as the state does not have adequate resources to provide universal education, private institutions are filling in the gap. The traditional madrassas, numbering 137 or fewer in 1947, had mushroomed to 10 000 by April 2002, with 1.7 million students, according to the Minister of Religious Affairs (Rahman, 2004b). Both madrassas and English-medium schools are eligible to receive government aid. Madrassas are funded and supported by private parties, donations, and religious organizations; the English-medium schools have the backing of the civil service and the armed forces. Most scientific, technical, and professional education in Pakistan follows the pattern observed in India, i.e., the medium of instruction is English.

The medium of instruction in Bangladesh is Bangla in most educational institutions; private institutions offer English-medium education to the children of the elite and Koranic education in Arabic to students of religion. Arabic is taught in all religious schools because Islam is the state religion. Members of religious minorities are exempt from learning Arabic or attending courses in religion.

The official policy in Nepal is to facilitate education through the mother tongue, and various programs are in place to implement that policy. According to the document put out by Ministry of Education and Sports (dated October 2004), Nepal will institute a national action plan in primary education between the years 2005 and 2007 that will translate a three-language policy into practice, the three languages being Nepali, a local language, and English.

The official policy in Sri Lanka is to encourage Sinhala and Tamil, the latter in its region, as mediums of education. English is not discouraged as a medium of instruction; it has been recognized as a medium for imparting higher education, as it is in India, Pakistan, and Bangladesh.

Role of English

Regardless of attitudes and perceptions, and of official policies encouraging mother tongues as mediums of instruction, demand for English education is on the rise all over South Asia, resulting in loss of this domain for other languages. For example, according to India's National Council of Educational Research and Training's Sixth All India Educational Survey (1998, based on a survey conducted in 1993–1994), the number of languages used as mediums of instruction in primary education declined from 47 in 1978 to 33, and in secondary education, from 43 in 1978 to 21.

Economic Development versus Issues of Identity

All countries of South Asia are struggling to find their niche in the global economy of the 21st century. The demands of development favor English, while the desire to preserve distinct cultural identities requires the use of national and often regional and ethnic languages in many domains of life. This is a struggle

in which the state sometimes comes to a decision that leads to political unrest, as in Sri Lanka. At other times, the state makes a decision, and the people are content to follow it, as in Nepal. In more complex multilingual situations, the state may make decisions, and the people may agitate and have the decisions modified in several ways. In India, following non-Hindi-speakers' agitations, the issue of official language was resolved by leaving English in that role for all practical purposes. Linguistic states were created for the same reason, and the list of national languages keeps expanding as more and more 'minor'-language speakers seek recognition and state support for their languages.

Trends

Except for Sri Lanka, countries in South Asia have had low rates of literacy, the lowest occurring in Bangladesh and Pakistan. The concerted efforts at providing compulsory primary education to children and spreading literacy among adult populations have produced impressive results in some countries (UNESCO Institute of Statistics 2000–2004). Female literacy is on the rise, and the numbers of schools and school-going children have grown several fold (Research, Reference, 2004; Rahman, 2004a). Literacy efforts conducted in the mother tongue or another familiar language have a much greater chance of success than those in an unfamiliar language. However, the increasing demand for English education at the primary level creates a dilemma. Providing support to multiple mother tongues and regional languages is an immense task; simultaneously imparting literacy and education in what to most school-going children is an unfamiliar language makes the task even more complex. In a global economy with increasing reliance on service-oriented occupations and emphasis on technical education for better-paying jobs, it is natural for parents – educated or not – to want English-medium education for their children.

All the countries of South Asia are attempting to make it possible to bring technology to all their peoples, irrespective of language. In several states of India, it has become possible for people in villages to log onto the Internet using their local languages to get information about water-usage patterns or marketability of their products at the highest margins of profit (www.apnic.net). Rahman (2004a, 2004c) pleads for extending 'localization' (i.e., translation of computer programs from English) to the minor languages of Pakistan. Such technical developments require considerable financial investment. Language education policies have the potential for realizing the goals of universal (technical) education; whether languages with small numbers of speakers will survive in competition with English and other major regional and official languages is difficult to predict.

See also: Bangladesh: Language Situation; Bilingual Education; Bilingualism; India: Language Situation; Language of Tertiary Education; Language Planning and Policy: Models; Language Policy in Multilingual Educational Contexts; Nepal: Language Situation; Pakistan: Language Situation; Sri Lanka: Language Situation.

Bibliography

His Majesty's Government, National Planning Commission Secretariat, Central Bureau of Statistics (2003). *Statistical yearbook of Nepal 2003.* Kathmandu.

Krishnamurti B (1995). 'Official language policies with special reference to the Eighth Schedule of the Constitution of India.' In Gupta R S, Abbi A & Aggarwal K N (eds.) *Language and the state: perspectives on the Eighth Schedule.* New Delhi: Creative Books. 8–23.

Krishnamurti B (1998). *Language, education and society.* New Delhi: Sage Publications.

Mallikarjun B (2001). 'Languages of India according to the 1991 census.' *Language in India* (monthly online journal). www.languageinindia.com, vol. 1(7).

Pakistan.gov. http://www.pakistan.gov.pk.

Pattanayak D P (1991). *Language, education and culture.* Mysore: Central Institute of Indian Languages. [CIIL Occasional Monographs Series, 46.]

Rahman T (2004a). Language policy and localization in Pakistan: proposal for a paradigmatic shift. Paper presented at SCALLA 2004, January 5–7, 2004, Kathmandu.

Rahman T (2004b). Polarized education. *Dawn* Internet Edition, 17 July 2004. www.dawn.com (accessed 10.27.2004).

Rahman T (2004c). *Denizens of alien world: a study of education, inequality and polarization in Pakistan.* Karachi: Oxford University Press.

Research, Reference and Training Division, Publications Division, Ministry of Information and Broadcasting, Government of India (ed.) (2004). *India 2004: a reference annual*, New Delhi.

Savada A M (ed.) (1993). *Nepal and Bhutan: country studies.* Washington, DC: Federal Research Division, Library of Congress.

Sharma B N (1985). *Medium of instruction in India: a backgrounder based on official documents of the government of India.* New Delhi: Central Secretariat Library, Department of Culture.

Language Education Policies in Southeast Asia

T Clayton, University of Kentucky, Lexington, KY, USA

Introduction

Contemporary Southeast Asia covers nearly 10 million square miles of mainland and island territory generally south of China, north of Australia, and east of the Indian subcontinent. From northwest to southeast, Southeast Asia comprises Burma, Thailand, Laos, Cambodia, Vietnam, Malaysia, the Philippines, Singapore, Indonesia, Brunei, and East Timor.

The 11 nations of Southeast Asia are immensely diverse linguistically. Among the approximately 550 million residents of the region can be found speakers of hundreds of native languages (Chong, Hmong, Javanese, for example), many national languages (Khmer [Central Khmer], Thai, Vietnamese, for instance), and several regional languages of wider communication (varieties of Chinese and Malay in particular). With the exception of Thailand, the countries of the region endured colonization and as a result experienced greater or lesser exposure to European languages – Dutch in Indonesia; English in the case of Brunei Darussalam, Burma, Malaysia, the Philippines, and Singapore; French for Cambodia, Laos, and Vietnam; and Portuguese in East Timor.

Despite the diversity within and among the nations of Southeast Asia, two general themes emerge in examining their language education policies. Notably, governments have used language policies in education in an attempt to forge postcolonial or contemporary identities at the national level: the valorization of one or more usually indigenous languages in education integrates with other policy initiatives intended to engender social cohesion in the multiethnic and multilingual nation-states. At the same time, all 11 nations of Southeast Asia have created a place for English in education, usually as a foreign language, for reasons that emerge in the following case studies. These brief studies describe government language education policies and, as such, do not necessarily represent the practices of private educational institutions.

Burma

Burmese is the language of textbooks and most instruction in primary, secondary, and tertiary schools in Burma. At secondary and university levels, mathematics and sciences classes are taught with English texts, though teachers generally lecture in Burmese. Instruction in non-Burmese indigenous languages is extremely unusual, though a few Mon-language schools have recently been established within the context of a separatist cease-fire agreement. All Burmese students study English as a foreign language at all levels of education. Students may select majors in many European and Asian languages at the University of Foreign Languages and other institutions. The University for National Races is the only educational institution in the country that teaches non-Burmese indigenous languages.

Contemporary education policies that support English language learning in Burma intend to renew global and international connections that were purposely severed by isolationist governments after independence (N. Cheesman, personal communication; see also Cheesman, 2003).

Thailand

The national language, Thai, serves as the medium of instruction in almost all government primary schools, secondary schools, and universities. Very recently, a few minority languages began to be taught as subjects in government schools; these languages include Chong, Lahu Shi, and Mon. All students study English as a foreign language starting in the fifth year of primary school. In university entrance examinations, English is treated as a subject with a value equal to others. In their third year of high school, students may select a third language from among Mandarin Chinese, French, German, and Japanese.

Foreign language policies supportive of English reflect the government's recognition of the value of this language for global business communication and for international education and travel (C. De'Ath, personal communication; see also Kosonen, 2003).

Laos

Lao, the national language, serves as the medium of instruction in almost all government schools. In some primary and secondary schools in minority regions, some teachers use local languages in the classroom, and some professors at some higher education institutions use French texts while lecturing in Lao. Students are required to study English as a foreign language in all schools beyond the primary level, and some primary schools in urban areas also teach English. Depending on the availability of teachers, students may additionally be able to study other foreign languages, among them Mandarin Chinese, French, and Japanese.

The emphasis on English as a foreign language in Laos anticipates communication with development agencies and participation in ASEAN (K. Kosonen, personal communication; see also Souvanvixay *et al.*, 2002).

Cambodia

Cambodian or Khmer (Central Khmer), the national language, serves as the medium of instruction in most primary and secondary schools, though a few bilingual initiatives (Khmer and various indigenous languages) have recently opened in primary schools in minority areas. Cambodians study either English or French as a foreign language in secondary schools in programs generally supported by development agencies; students tend to choose English over French by a two-to-one margin. Most students in public universities study in Khmer and learn English as a foreign language, though some enrolled in programs supported by the French government or the Agence Universitaire de la Francophonie receive French-medium instruction.

Cambodia's dual foreign-language policy reflects the country's participation in diverse, and sometimes competing, international networks that privilege the two languages, notably ASEAN and the Organisation Internationale de la Francophonie (see Clayton, 2002).

Vietnam

Vietnamese is the language of education in schools of all levels in Vietnam. Some minority languages, including Cham, Hmong, and Khmer (Central Khmer), are taught as subjects in some primary schools. Students choose among Mandarin Chinese, English, French, and Russian to satisfy the foreign-language requirements that begin in primary education. Not all schools offer all four foreign languages; approximately 90% offer English, with successively fewer teaching French, Mandarin Chinese, and Russian. English as a foreign language is a *de facto* requirement in secondary education, however, for students must sit for an English language test as part of their secondary school-leaving examinations.

The emphasis on English in Vietnamese schools reflects the government's desire to participate in diverse economic and political venues and organizations that use this language (G. Fry, personal communication; Pham L. H., personal communication; see also Wright, 2002).

Malaysia

In primary schools in Malaysia, students may study in Bahasa Malaysia (a variety of Malay), Mandarin Chinese, or Tamil. Until recently, Bahasa Malaysia served as the sole language of instruction in secondary schools and universities, and students studied English as a foreign language at all levels of education. In 2000, public universities began offering between 30% and 50% of their courses in English, and in 2003 mathematics and science classes began to be taught in English in primary and secondary schools. In 1999, university applicants began to be required to take the Malaysian University English Test to assess their English proficiency.

The government's recent movement toward English as the language of the global economy and global information flows reverses a trend away from this language after independence (M. Lee, personal communication; see also David and Govindasamy, 2003).

The Philippines

In primary schools, Filipino, the Tagalog-based national language, serves as the medium of instruction in all subjects except mathematics and sciences, which are taught in English. The balance shifts in secondary schools, where all subjects except Filipino language and social studies are taught in English. Most instruction at Philippine universities is also given in English, except in Philippine history and government courses, where professors teach in Filipino. All students study English as a foreign language at all levels of education. At universities, students may elect to study other foreign languages, including French, Japanese, and Spanish.

The increasing emphasis on English as a foreign and instructional language anticipates participation in regional and global venues where English serves as the language of communication, among them the overseas employment market for Filipino professionals (A. Gonzalez, personal communication; see also Hidalgo, 1998).

Singapore

English serves as the medium of instruction at all levels of government education in Singapore, with the exception of civics and morals classes in primary and some secondary schools. Students are expected to choose one of the state's other official languages – Mandarin Chinese, Malay, or Tamil – to study as a 'mother tongue language,' and they are tested on their knowledge of this subject in school-leaving examinations. Students from north Indian ethnic backgrounds may fulfill this latter requirement with Bengali, Gujarati, Hindi, Punjabi (Eastern Panjabi), or Urdu, which they study in nonformal classes organized by their respective language communities.

The choice of English as the language of instruction advances two objectives: it provides a medium for neutral, intranational communication while at the same time preparing graduates for economic and other global and international interactions (J. Tan, personal communication; see also Gopinathan *et al.*, 1998).

Indonesia

The national language, Bahasa Indonesia (Indonesian), is the language of virtually all instruction at all levels of government education in Indonesia, and this variety of Malay also provides the language for all textbooks. The national curriculum dictates that all Indonesians study English as a foreign language in secondary schools, and the school-leaving national achievement test includes a section on English language skills. Additionally, different local or regional languages such as Batak or Javanese may be taught as subjects in 'local content' classes in primary and secondary schools. In tourism-dependent regions such as Bali, English may be introduced in local content courses.

The Indonesian government emphasizes English as a foreign language as a means of facilitating involvement in international forums that use this language, notably the Association of Southeast Asian Nations (ASEAN) and the World Trade Organization (C. Bjork, personal communication; B. Leigh, personal communication; see also Kuipers, 1998).

Brunei Darussalam

In Brunei, children study in Malay until the age of eight, at which point they transfer to English-language instruction in geography, history, and mathematics classes. English serves as the language of all secondary education, except in those subjects (Malay literature or Malay Islamic monarchy for instance) that have a clearly Malay or Islamic orientation. English continues to be used virtually exclusively at the university level. English is taught as a foreign language in all the state schools, beginning in primary education. Public schools funded by the ethnic Chinese community additionally teach Mandarin Chinese as a foreign language.

The government of Brunei intends its English-as-a-foreign-language policy to prepare bilinguals capable of economic communication within the region and, particularly, with the United Kingdom (S. Upex, personal communication; see also Gunn, 1997).

East Timor

East Timor's two official languages, Portuguese and Tetum, serve as dominant media of education in government schools. At the primary level, the availablility of teaching materials makes it possible to teach both languages from the first year. In secondary schools, many teachers use readily available Portuguese books while providing classroom explanation in Tetum; some teachers continue to use Bahasa Indonesia (Indonesian), though this practice is to be phased out by 2006. A similar mix of educational media prevails in universities, where in addition some foreign lecturers give classes exclusively in Portuguese and English. Both English and Malay are taught as second languages in East Timorese high schools, with English being the more popular choice among students.

The East Timorese government expects knowledge of English to complement and supplement that of Portuguese in regional and international communication (G. Hull, personal communication; see also Appleby, 2002).

See also: Brunei Darussalam: Language Situation; Burma: Language Situation; Burmese; Cambodia: Language Situation; English in the Present Day (since ca. 1900); Indonesia: Language Situation; Khmer; Lao; Laos: Language Situation; Malay; Malaysia: Language Situation; Philippines: Language Situation; Singapore: Language Situation; Southeast Asia as a Linguistic Area; Tagalog; Thailand: Language Situation; Vietnam: Language Situation; Vietnamese.

Bibliography

Appleby R (2002). 'English and East Timor.' In Lo Bianco J (ed.) *Development and language: global influences and local effects.* Melbourne: Language Australia. 23–35.

Cheesman N (2003). 'School, state and Sangha in Burma.' *Comparative Education 39,* 45–63.

Clayton T (2002). 'Language choice in a nation under transition: the struggle between English and French in Cambodia.' *Language Policy 1,* 3–25.

David M K & Govindasamy S (2003). 'Language education and "nation building" in multilingual Malaysia.' In Bourne J & Reid E (eds.) *Language education.* London: Kogan Page. 215–225.

Gopinathan S, Pakir A & Ho W K (eds.) (1998). *Language, society and education in Singapore: issues and trends* (2nd edn.). Singapore: Times Academic Press.

Gunn G (1997). *Language, power, and ideology in Brunei Darussalam.* Athens, OH: Ohio University Center for International Studies.

Hidalgo C A (1998). 'Language choice in a multilingual society: the case of the Philippines.' *International Journal of the Sociology of Language 130,* 23–33.

Kosonen K (2003). 'Community participation in minority language education in Thailand.' *Journal of Southeast Asian Education 4(1),* 104–136.

Kuipers J C (1998). *Language, identity, and marginality in Indonesia.* Cambridge: Cambridge University Press.

Souvanvixay M, Nouannavong O & Keovongsa K (2002). 'Second languages in ethnic and linguistic diversity in Laos.' In Lo Bianco J (ed.) *Development and language: global influences and local effects.* Melbourne: Language Australia. 295–302.

Wright S (2002). 'Language education and foreign relations in Vietnam.' In Tollefson J (ed.) *Language policies in education.* Mahwah, NJ: Lawrence Erlbaum. 225–244.

Language Education Policies in the Pacific

R Baldauf and W Cheung, The University of Queensland, St Lucia, Australia

Introduction

The Pacific region, comprising about 25 000 islands, is traditionally divided into three areas: Melanesia, Micronesia, and Polynesia. To this can be added three larger states, Australia, Indonesia, and New Zealand, which together comprise the 29 polities of the Pacific that provide the focus of this article (see Table 1).

Language Situation in the Pacific Polities

The Pacific is characterized by a wide variety of geographic and demographic features (Table 1). The largest geographic polity, Australia (some 7 million square kilometers), is 1.2 million times bigger than the smallest one, Wake Island (6.5 square kilometers). Indonesia, with about 238 million people, is the most populous polity, whereas Pitcairn Island has only 46 inhabitants. These great differences among the polities are also reflected in the diversity of the region's languages and cultures. The Pacific is

Table 1 Pacific Polities – Demographic and Geographic Information

Polity	*Population*[a]	*Area (sq. km)*	*Population age between 0–14*	*Adult Literacy Rates %*
Melanesia				
New Caledonia and Dependencies[c]	213 679	19 060	29.4%	91.0
Papua New Guinea[j]	5 420 280	462 840	38.3%	44.9
Republic of the Fiji Islands[j]	880 874	18 270	31.7%	92.9
Republic of Vanuatu[j]	202 609	12 200	34.1%	33.0
Solomon Islands[j]	523 617	28 450	42.4%	20.5
Timor-Leste (East Timor)[j]	1 019 252	15 007	37.8%	43.0
Micronesia				
Federated States of Micronesia[j]	108 155	702	37.7%	90.0
Guam[d]	166 090	549	29.8%	96.4
Northern Mariana Islands[e]	78 252	477	20.3%	99.0
Republic of the Marshall Islands[j]	57 738	181.3	38.6%	90.8
Republic of Nauru[j]	12 809	21	38.2%	99.0
Republic of Palau[j]	20 016	458	26.6%	78.0
Wake Island[d]	200[b]	6.5	NA	NA
Polynesia				
American Samoa[d]	57 902	199	36.6%	99.5
Cook Islands[f]	21 200	240	NA	100.0
French Polynesia[c]	266 339	4 167	27.5%	97.8
Kingdom of Tonga[j]	110 237	748	37.1%	95.9
Niue[f]	2 156	260	NA	NA
Norfolk Island[g]	1 841	34.6	20.2%	NA
Pitcairn Island[h]	46	47	NA	NA
Republic of Kiribati[j]	100 798	811	39.3%	92.2
Samoa[j]	177 714	2 944	28.3%	98.2
State of Hawai'i, USA	1 211 537	16 649	27.1% (under 19)	NA
Tokelau Islands[i]	1 405	10	42%	91.2
Tuvalu[j]	11 468	26	31.3%	99.5
Wallis and Futuna[c]	15 880	274	NA	95.0
Larger States				
Australia[j]	19 913 144	7 686 850	20.1%	NA
New Zealand[j]	3 993 817	268 680	21.7%	99.5
Republic of Indonesia[j]	238 452 952	1 919 440	29.4%	88.5

[a] Estimated in July 2004.
[b] No indigenous inhabitants.
[c] French territory.
[d] United States unincorporated territory.
[e] Commonwealth in political union with the U.S.
[f] Self-governing, in free association with New Zealand.
[g] Australian territory.
[h] UK territory.
[i] New Zealand territory.
[j] Independent States.

a high-density language and ethnic group region that encompasses between 20 and 30% of the world's languages, many of which are increasingly endangered.

Except for Wake Island, which has no native inhabitants, all the other polities are multilingual (see Table 2). Papua New Guinea has the largest variety of languages with about 832, but Australia, Indonesia, the Solomon Islands, and Vanuatu are polities in which 85 or more languages are used. In most polities a number of languages are used for communication purposes, but not all the dialects, pidgins, or vernaculars that exist are well documented, and the number of users of some of these language varieties may be limited to a small proportion of the total population. For instance, in New Caledonia, inhabitants use 33 Melanesian-Polynesian varieties, but only 17 of them are spoken by more than 1000 people. This complex language situation provides the conditions for lingua franca to develop for everyday communication. Although Melanesian pidgin is the common lingua franca in that region of the Pacific, for historical reasons, English is the common lingua franca across the Pacific. English has national language status in 16 of the polities and official language status in seven others, while it is the first foreign language in the rest of the Pacific (i.e., in Timor Leste (Portuguese), Indonesia (Bahasa Indonesia), French Polynesia, New Caledonia, and Wallis and Futuna). In fact, a lingua franca, rather than a vernacular, is often chosen as an official language or national language in Pacific polities, such as Bahasa Indonesian in Indonesia.

This complex multiethnic language situation creates both risks and opportunities for language planning in each polity. Subtle language planning is needed to overcome the problems such complex language situations add to economic, social, and political development as well as to everyday life. Governing authorities could impose their favored language solutions on the populace, raising the status of some languages and reducing that of others, but there are risks in doing so. Language represents the culture of its users, and being excluded from official or national status means being subjugated politically, socially, or even economically, leading to possible cultural conflicts within or across national boundaries. Hence, 'keeping balance' and making symbolic statements becomes important. Thus, Maori and English are the official languages of New Zealand, and Portuguese and Tetum, not Indonesian, were selected for official and national use when East Timor became independent.

Language-in-Education Planning in the Pacific Polities

A great diversity of educational conditions and provision exists across the polities in the Pacific. As Table 1 indicates, some polities have high school age populations – and consequent costs – with proportions of school-aged students ranging from 42.4% in Solomon Islands to 20.1% in Australia. Educational outcomes, as measured by adult literacy rates, also vary widely, particularly in Melanesia.

In a number of polities, education is linked with national development. For instance, in Fiji, education policies are formulated within the framework of national goals so that the process of education contributes also to the achievement of government's broad outcomes. However, some Pacific polity governments are unable to support full educational provision and depend to varying degrees on support from the larger states (see Table 3). One outcome of this educational dependence, language-in-education planning in Pacific polities, has a focus on English, and to a lesser extent, French, Portuguese, or Bahasa Indonesian. A number of specific reasons for educational dependence can be specified.

First, in some larger states, formal education systems have been developed locally (e.g., Australia, Indonesia, or New Zealand), but many of the others are modeled on or dependent on those of the (former) colonial powers (e.g., Kiribati on the British, American Samoa on the United States, Papua New Guinea on Australia) due to their small population and education markets. As these small Pacific polities have difficulties in developing their own teaching materials, they import them and sometimes expatriates to help with the teaching. Furthermore, educational assessments in these small polities are modeled on, if not copied from, systems in the larger states (e.g., English language assessment in the Cook Islands or the school system in Kiribati draws heavily from New Zealand). Nearly 40% of Cook Island students continue on to Form 5 to complete the New Zealand School Certificate. Although Fiji is only now developing its own assessment system, much of the Micronesia relies on U.S. standardized tests. This kind of educational dependence is funded through aid provided by larger polities, and these dependencies tend to strengthen English or other exogenous languages, such as French, at the expense of indigenous languages.

Second, the lack of higher education institutes in these small polities also has reinforced the importance of English or French in language-in-education planning. Although the University of the South Pacific

Table 2 The Language situation in the Pacific Polities

Polity	*Languages spoken*	*Role of English*	*Role of other languages*
American Samoa	English, Samoan	Official language	Samoan: official language, spoken by 90% of the population.
Australia	English, 200 migrant and 90 indigenous languages	National language	Arabic, Mandarin, French, German, Indonesian, Italian, Japanese, Modern Greek, Spanish, Korean, Russian, Thai, Vietnamese, and some Aboriginal languages are taught in schools are LOTE.
Cook Island	English, Maori, Pukapukan	National language	Maori: 6 different dialects, not all have written forms. Pukapukan: Not intelligible with other Cook Islands languages.
Timor Leste	English, Portuguese, Tetum, Indonesian	Unofficial utility language of UN, peacekeepers and aid agencies	Portuguese: official language. Tetum: national language. Indonesian: an unofficial utility.
Fiji	English, Bau Fijian, Tamil, Hindustani, Rotuman	Official language	Bau Fijian: used as the standard Fijian along with various dialects, also the traditional lingua franca among Fijians. Hindustani and Tamil: spoken by Indo–Fijians.
French Polynesia	French, Tahitian	Foreign language	French: national language. Tahitian: community language.
Guam	English, Chamorro, Japanese	National language	Chamorro: spoken by about 50% of population. Japanese: used in tourist industry.
Hawai'i	English, Hawaiian	Official language	Hawaiian: spoken on all of the inhabited islands of Hawaii, reestablished as an official language of the state in 1978.
Indonesia	Bahasa Indonesia, Javanese, Dutch, Sundanese, 400+ other languages	1st FL in secondary schools. A tertiary level 'library language'	Bahasa Indonesian: official language, major L2. Javanese: most widely spoken (40%) variety. Sundanese: 15% of the population spoken. Dutch: former colonial language.
Kiribati	English, Kiribati,	National language	Kiribati: national language spoken by 97% of the population.
Marshall Islands	English, Marshallese	Official language	Marshallese: widely spoken as a second official language.
Federated States of Micronesia	English, Chuukese, Pohnpeian, Kosrean, Yapese, Ulithian, Kapingamarangi, Woleaian, Nukuo	National language	Chuukese: also known as Trukese, the most widely spoken indigenous language, followed by Pohnpeian. Other indigenous languages: All have more than 10 000 speakers.
Nauru	English, Nauruan	National language	Nauruan: limited intergenerational transmission, still actively used in the Republic, used as a second language by the Kiribati, Tuvalu, Kosraean, and Marshallese.
New Caledonia	French, 33 Melanesian–Polynesian dialects	Foreign language. English signage in tourist-related government departments	French: national language also is the first language of 35.6% of the population. The three most popular dialects: Dehu (15 000 speakers), Wallisian (9000), New Caledonian Javanese (6750).
New Zealand	English, Maori	National language	Maori: official language. Many migrant languages.
Niue	English, Niuean	National language	Niuean: spoken by 97.4% of the population.
Norfolk Island	English, Pitcairn–Norfolk	National language	Pitcairn–Norfolk: an in-group language used to assist in the preservation of identity, second language only.
Northern Mariana Islands	English, Chamorro, Carolinian	National language	Chamorro: national language, gaining its importance, is taught at the University of Guam. Carolinian: national language, semiofficial status.
Palau	English, Palauan, Sonsoralese, Tobi, Angaur; Japanese	National language	Palauan: national language in all states except Sonsoral (where people speak Sonsoralese), Tobi (Tobi), and Angaur (Angaur, Japanese).

Continued

Table 2 Continued

Polity	*Languages spoken*	*Role of English*	*Role of other languages*
Papua New Guinea	English, Tok Pisin, Hiri Motu, 832 indigenous language	Official language	Tok Pisin: official language, main lingua franca, most frequently used language in Parliament and commerce. Hiri Motu: official language, a pidginized Motu, influenced by English, Tok Pisin, and Polynesian languages, but unintelligible to Motu speakers.
Pitcairn Island	English, Pitcairnese, Pitcairn–Norfolk	National language	Pitcairnese: mixture of an 18th-century English and Tahitian dialects used in informal conversation among the Pitcairners.
Samoa	English, Samoan	Official language	Samoan: official language, spoken by 93% of the population
Solomon Islands	English, 60 indigenous languages, 120 Melanesian Pidgins	National language	Melanesian Pidgin: popular lingua franca with no official status, rarely used in official contexts. Twelve indigenous languages have more than 10 000 speakers, largest is Kwara'ae with 33 600 speakers.
Tokelau	English, Tokelauan	National language	Tokelauan: used in the schools.
Tonga	English, Tongan	Official language	Tongan: national language.
Tuvalu	Tuvaluan, Kiribati	Foreign language	Tuvaluan: official language. Kiribati: used on the island of Nui, part of Tuvalu culture.
Vanuatu	English, French, Bislama. About 87 indigenous languages	National language	French: national language. Bislama: national language, widely used in commerce, government, and internal dealings.
Wake Island	English	National language	U.S. territory. No indigenous residents.
Wallis and Futuna	French, East Futuna, Wallisian	Foreign language	French: national language. East Futuna: used in east Wallisian and Futuna. Wallisian: used by 7500 in Wallis and Futuna

and the University of Guam meet some needs of the smaller states, most polities are dependent on educational opportunities available from Australia, New Zealand, or the United States. For example, almost all current government leaders in the FSM are graduates of U.S. colleges. Thus, learning the languages of the (former) colonial powers is important for seeking further education opportunities in the region.

Such educational dependence secures English, French, or Indonesian's dominant language status in these polities, making it more difficult for indigenous or migrant languages to find instructional space. Consequently, language-in-education planning and language education strengthen and stabilize the dominant position of exogenous languages.

Although the pressures to learn English (or French or Indonesian) are great, some of the polities are trying to take account of the evidence that shows that education in the first language, rather than in the former colonial language, has educational advantages and closely associated psychological, social, and cultural benefits. Hence, some polities use their vernacular languages for at least part of formal education. An outstanding example of this is Tonga, which has high literacy rates for both Tongan and English. In PNG, which previously emphasized only English, a grassroots nonformal education movement arose in the 1980s that insisted on the use of vernaculars for elementary education to stem the tide of massive educational failure. For most of the polities, vernacular languages are only used as a medium of instruction for the very early years of schooling (e.g., in Cook Islands, Guam, Samoa, or Indonesia) and are then replaced by English or other lingua franca, although in some (e.g., American Samoa) indigenous language and culture continues to be taught as a subject.

In the region, official languages or national languages are subjects student teachers are required to study during their training. Some Pacific polities integrate the pedagogy of official languages or national languages in their teacher training programs (e.g., Fiji, Kiribati, or Samoa); some just indicate the language proficiency requirements for these programs. In addition, it is quite common to have examination requirements, at different levels, to assess the students' language proficiency. Some polities introduce such examinations at the primary level (e.g., Cook Islands, Guam, or Indonesia), some at secondary (e.g., Fiji, Kiribati, or Solomon Islands), but most

Table 3 Language-in-Education planning in the Pacific Polities

Polity	*Education system*	*Language-in-Education planning*
American Samoa	English is taught in public elementary, middle, and high schools. All content areas are taught in English. Samoan Language is included in the K–12 curriculum.	
Australia	School education is compulsory until the age of 15. Tasmania up to 16.	Language of instruction is English. A few schools offer bilingual or non-English immersion programs. Indigenous languages may also be used in primary education in some regions. LOTE is commonly taught as a subject in public schools. Formal language planning is the work of state governments and the federal government. For example, the National Language Policy Documents by Commonwealth Government in the 1980s and 1990s.
Cook Islands	Compulsory and free from the age of 6 to 15.6 years primary and 6 years secondary. Many students continue their education in New Zealand or Fiji. Use New Zealand Qualifications Authority standard for assessment. Cook Islands Assessment Board determines curricular design. English and Maori are the compulsory subjects of the Cook Islands School Certificate Examination. English is the compulsory subject of the Six Form Certificate.	The medium of instruction in primary schools may be English or Maori, but where Maori is used, particularly in the North Group of islands, English often replaces it after the first two years of schooling. The medium of instruction in secondary schools is English.
Timor Leste	Rehabilitation of education facilities is being undertaken. High levels of illiteracy especially among women. Too few schools for large numbers of school-age children. Cost of schooling means many children, especially girls and orphans, do not attend. Not enough qualified teachers. Very large class sizes. Schools lack furniture, equipment, textbooks, and supplies.	In 1975, Indonesian became the language of the state and the medium of education. Portuguese became the language of the resistance to Indonesian rule. Tetum was strengthened through its association with the Catholic church. United Nations Peace keeping Force, international aid agency workers, and UNTAFT raise the demand for English to gain access to economic, social, and political benefits of the foreigner influx.
Fiji	Education is not currently compulsory throughout Fiji. Compulsory education was only introduced in two of the nine school districts on a trial basis. Most of the schools are communal and essentially monoracial, i.e., run by Fijian or Indo–Fijian local committees and religious organizations, though government policy is to see schools as integrated as possible. In 2004, there were 712 primary schools. Of these only 20 are actually run directly by the government. A 6-year program and 8-year program are available for primary education. Secondary education consists of a 7-year program (Form 1–7). But few of the secondary schools provide Form 7. Secondary schooling is neither free nor compulsory.	English is the main language of instruction throughout the education system, including all secondary and higher education, and is the compulsory core subject for the Fiji Junior Certificate Examination and Fiji School Leaving Certificate. Primary schools may use a vernacular language (most commonly Fijian or Hindi) for the first three years (grades 1–3), while urban schools (with more mixed enrolments) generally use English from grade 1. From grade 4 all schools are required to use English as the language of instruction, with vernacular languages being offered as subjects. Students study their mother tongue throughout primary and secondary education, and Fijian and Hindi may also be studied as part of undergraduate programs at the University of the South Pacific, up to second-year level. English plus social science is one of four streams prospective teachers can choose from their training. Since 1987 a program has been initiated to teach Hindi and Fijian cross-culturally in primary schools. English is rapidly replacing vernacular bilingualism and the use of local contact languages (Pidgin Fijian and Pidgin Hindustani) as the main means of communication, especially in urban areas. In the government, English is designated as the official language of Parliament, Fijian and Hindustani are only occasionally used in the Senate and hardly ever in the House of Representatives.
Guam	Compulsory from age 5 to 16. Every school year during the 4th quarter, students in grades 1, 3, and 5 take the SAT9 test, which includes areas of reading, language, and listening.	The Department of Education provides courses for students with first languages other than English. Students are screened based on the language spoken at home. If a second language is spoken at home, a student may be tested and enrolled in appropriate English classes.

Continued

Table 3 Continued

Polity	*Education system*	*Language-in-Education planning*
Hawai'i	Total of 279 public schools follow the U.S. system, where children attend preschool, then elementary school grades K–5, intermediate or middle school grades 6–8, high school grades 9–12. Universities cater to local students as well as those from the Pacific.	In 1987, government schools began using Hawaiian as the medium of instruction at selected sites. In 1990, the federal government of the United States adopted a policy to recognize the right of Hawaii to preserve, use, and support its indigenous language. Hawaiian is the only Native American language that is used officially by a state government.
Indonesia	Secular and Islamic education is offered at every level of the education system, 6 years primary and 6 years secondary. Since 1995, compulsory education has been extended to 9 years. Schools are supervised by the Ministry of Education and Culture (MOEC), or by the Ministry of Religious Affairs for Islamic schools. Centralized curricular design. The role of the private sector in all levels of education is substantial. Language Council coordinates language issues, particularly for new scientific and technical vocabulary development. In grade 12 of senior secondary school, there is the National Evaluation of Studies at the Final Level – Academic Senior Secondary School, which includes English and Bahasa Indonesian.	The medium of instruction is Bahasa Indonesia, as stipulated by law. The local regional language may be used during the first 3 years of primary education in parts of Indonesia that have major regional languages (e.g., Javanese, Sundanese, and Balinese). The teaching and use of Chinese is discouraged. Dutch has virtually disappeared in public schools. The compulsory study of English is usually begun at junior secondary school.
Kiribati	7 years of primary, 6 years of secondary, and 1 year of preuniversity studies (Form 7). Secondary education, which is neither free nor compulsory, offers limited places, but leads to possible entry into the University of South Pacific or other universities. Staffing consists mostly of Kiribati graduate teachers but expatriate positions have been supported by aid for many years. The Curriculum Development and Research Centre is in charge the curricular design. National Common Entrance Examination for primary students includes the test of English and Kiribati. In both the Kiribati Junior School Certification Examination and the Kiribati National Certificate Examination, English is nationally set and marked.	A rural Teaching English as a Second Language (TESL) program is continuing to expand. All lessons are in English. Admission to secondary school is based on examination – a major component of which is mastery of English. Language use in Kiribati is fairly uniform; no minority groups that might affect educational policy. English is included in the 3-year primary teacher-training program at Tarawa Teachers College.
Marshall Islands	Compulsory education for ages 6 through 14 or completion of eighth grade. A high school entrance examination is administered to all eighth graders to determine the approximately 300 students who will be admitted to the two public high schools each year.	
FSM	All FSM children are required to attend school through the grade 8 (age 5–14). Private educational facilities play an important role in the overall education system. Postsecondary, FSM students can go to the College of Micronesia, the University of Guam or other U.S. colleges.	
Nauru	Free and compulsory education, age 5–16. There are 6 infant schools, 2 primary schools, a secondary school, and a technical school; also a mission school. Scholarships are available for higher education overseas.	
New Zealand	Compulsory education is divided into primary, intermediate, and secondary schooling. Primary schools cater for children from the age of 5 (year 0 to year 6). Children in year 7 and 8 may either be in a separate intermediate school or part of a primary, secondary, or composite/area school. Secondary schools usually provide for students from year 9 to year 13. In addition to the mainstream schooling, students can choose *kure kaupapa*, Maori state-supported schools, where the teaching is in the Maori language and is based on Maori culture and values. The national curriculum is applied.	Maori could be used by Members of Parliament and used in at least one Court, the Maori Land Court. However, English/Maori diglossia has been shifting toward, English monolingualism since the end of the WW II. In 1985, primary schools that use Maori as instruction medium were established. In these schools, the whole curriculum was delivered by using Maori. There are Maori language national curriculum statements for Maori language as well as for other learning areas. Other languages have not obtained similar supports from the government.

Table 3 Continued

Polity	*Education system*	*Language-in-Education planning*
PNG	There are three levels of education administration, namely the central, provincial, and district. Nearly 3000 primary and secondary schools are there in PNG. 6 universities provide higher education in there. Language and Literacy Section in National Department of Education sets up the National Literacy and Awareness Council to improve the literary rate in the country.	English-only education had been implemented during the immediate years of PNG's independence. Since 1986, the government, with the help of overseas aid donors, has embarked on a massive reform of the English-only formal education system. The first 3 years of formal education will be in vernacular language at elementary school, run by local communities. Changes in educational policy with regard to the use of vernaculars in the formal sector began in 1991.
Pitcairn Island	There are 5 students on the island. Age from 6–15. And there is one 'school custodian and culture teacher.' No formal education system.	
Samoa	Primary schooling is compulsory. Eight years primary and 5 years secondary. Staffing is a problem in both government and mission secondary schools. Large percentage of the staff are expatriates, on contract or volunteers from the U.S. Australia, and New Zealand. Only about a third of those who have finished teacher education program remain in the country. One-quarter of teachers in Catholic secondary schools have no teaching qualifications. In the year 8 National Examination, English and Samoan are examined with papers set and marked locally. English and Samoan are compulsory subjects for the year 11 National Examination. In the Western Samoa School Certificate Examination, only English is the compulsory subject (a new assessment mechanism is under discussion).	The official medium of instruction is English, although Samoan may be used where appropriate, particularly in junior levels, to clarify difficult concepts. Samoan language has been integrated into the curriculum at all levels. The Samoan language is used for conducting proceedings in the parliament with simultaneous translation into English. English is also used in the business and commercial sectors and is widely spoken. English and Samoan languages are included in a 3-year teacher training diploma program.
Solomon Islands	Education is compulsory from the age of 6 to 15; 6 years primary education and 5 years secondary, the first 3 years of which takes place in provincial secondary schools and in community high schools. Completion of form 3 represents the end of school education for most students. Current educational practices in the country are a continuation of colonial practices, and there continues to be a divide between the better-resourced schools in the regional centers and those on the rural areas. Of all secondary teachers 29% are expatriates. Major shortages of qualified teachers in English. English is included in both the Solomon Islands Secondary Entrance Examination and Solomon Islands Form 3 Examination.	English is the sole medium of instruction despite Melanesian pidgin being the common lingua franca. English is the language of instruction in the University of South Pacific. No natural Melanesian English-speaking community having developed. Although English is only spoken by 10–15% of Solomon Islanders, it is a requirement for all midlevel or higher jobs and for upward mobility in the government or private sector. English is included in the 3-year program of secondary teacher training.
Tonga	Free education is provided for children aged between 6–14. The 6 years of primary education are compulsory. Secondary schooling is free to the age of 14 but not compulsory. Churches and other private organizations play important roles complementary to the government's role in education. There are 26 secondary schools but only the government administrates six. Secondary Entrance Examination – both English and Tongan are subjects of examination.	The kingdom has high level of literacy, particularly in the vernacular language. Written Tongan is used for religious, governmental, informational, commercial, and personal functions. English is for external relations, higher-level administration, and education. Although the educational system has encouraged the learning of English, it has maintained an emphasis on the Tongan language. Tongan is the medium of instruction in primary school, whereas English is a second language. At the secondary level, English is the medium of instruction, although teachers have the option to use Tongan in Forms 1 and 2. Both English and Tongan are core studies of secondary teacher diploma program. Advanced English language course is offered.
Tuvalu	The only secondary school for Tuvalu islanders is located at Vaitupu.	The Vaitupu dialect is the dominant and most widely spoken and is the dialect of the educated, as it was used by colonial radio broadcasters and for the former weekly Tuvalu language newspaper.

Continued

Table 3 Continued

Polity	*Education system*	*Language-in-Education planning*
Vanuatu	The years of schooling of primary, junior secondary, and senior secondary are 6, 4, and 3–4, respectively. High primary enrolment rate but low in secondary. Although there were 284 primary schools run by the government, only 58 government secondary schools (in 2000). There is a campus of the University of the South Pacific set up in the capital.	The principal languages of education are English and French, but under revised education policies, vernacular teaching is being introduced in some rural schools.

test at both levels. Although these examinations may not cover only language assessment, failure on these language examinations in some polities may mean the lost of further education opportunities (see Table 3).

Future Language Policies and Education Trends

The language and language-in-education planning situation in the Pacific both influence and are influenced by current language education policies that expand the roles of English, resulting in some common outcomes. English is increasingly being used by English-knowing and English-using bilinguals, and varieties of English are beginning to assert themselves as legitimate ways of speaking. For pragmatic reasons, English continues to strengthen in this region, with the number of first language speakers increasing as more of the elites become English-dominant bilinguals. This pragmatic use of English and its presence via the media creates a large group of English-conscious bilinguals, who can understand the language in limited contexts. Promotional policies, educational aid relationships, and practical benefits make English the language that is needed for advancement in education and white-collar employment. In urban areas English is beginning to infiltrate the home, which has been the primary domain of vernacular language use.

However, there are strong traditions of both multilingualism and using lingua francas as buffers that have allowed local languages to survive and prosper, serving as markers of identity. Will English continue to remain primarily a lingua franca with indigenous languages providing the primary source of identity, or will its increased adoption lead to the collapse of the linguistic ecology and the extinction of most indigenous languages? This situation is complicated by the development and adoption of local standardized languages as lingua franca within Pacific polities. Will this growing layer of folk plurilingualism primarily serve to strengthen the role of the first- and second-tier languages, thereby contributing to unbalanced diglossia, a shift away from the use of smaller vernacular languages, and reduced natural language variation?

Plurilingualism is likely to remain the norm in the Pacific, but the degree of linguistic variation may be significantly reduced unless proactive measures are taken. These language problems and potential solutions are bound up with broader political, cultural, and economic issues. But meeting national and international communicative needs, rather than saving languages from extinction, is likely to be the focus of language education policies and what they can deliver.

See also: Bilingual Education; Education in a Former Colonial Language; Education in a Multilingual Society; Language Education for Endangered Languages; Language Education for Language Revival; Language Education Policies in North America; Language Education Policies in South Asia; Language Education Policies in Southeast Asia; Minority Languages: Education; Mother Tongue Education: Nonstandard Language.

Bibliography

Appleby R (2004). 'The political context of English language teaching in East Timor.' In Hickling-Hudson A, Matthews J & Woods A (eds.) *Disrupting preconceptions: postcolonialism and education.* Flaxton, QLD: Post Pressed. 235–249.

Appleby R, Copley K, Sithirajvongsa S & Pennycook A (2002). 'Language in development constrained: three contexts.' *TESOL Quarterly 36(3),* 323–346.

Baldauf R B Jr (1990). 'Education and language planning in the Samoas.' In Baldauf R B Jr & Luke A (eds.) *Language planning and education in Australasia and the South Pacific.* Clevedon: Multilingual Matters. 259–276.

Baldauf R B Jr & Djité P G (2003). 'Australasia and the South Pacific.' In Maurais J & Morris M A (eds.) *Languages in a globalising world.* Cambridge: Cambridge University Press. 217–227.

Benton R A (1981). *The flight of the Amokura: Oceanic languages and formal education in the South Pacific.* Wellington: New Zealand Council for Educational Research.

CIA, The Government of the United States of America (2003). *The world factbook 2003.* http://www.cia.gov. Accessed on 6 April 2004.

Crowley T (2000). 'The language situation in Vanuatu.' *Current Issues in Language Planning 1,* 47–132.

Fuchs E M (1998). 'Englisch-tuvaluanische interferenz und sprach(en)politik in Tuvalu' [English-Tuvaluan Interference and Language Policy in Tuvalu]. *Asian and African Studies 7(1),* 10–22.

Gegeo D W & Watson-Gegeo K A (2001). 'The critical villager: transforming language and education in Solomon Islands.' In Tollefson J W (ed.) *Language policies in education: critical issues.* Mahwah, NJ: Lawrence Erlbaum. 309–325.

Grimes B F & Grimes J E (eds.) (2000). *Ethnologue: languages of the world* (14th edn.). Dallas, TX: Summer Institute of Linguistics <online version>. http://www.ethnologue. com. Accessed on 7th April 2004.

Gannicott K G (1990). 'Education in Fiji: the efficiency-equity quandary once again.' In Gannicott G K (ed.) *Pacific policy paper 6: education for economic development in the South Pacific.* Canberra: National Centre for Development Studies, The Australian National University.

Government of East Timor (2002). *Fact sheet 14.* http://www.gov-rdtl.org. Accessed on 6th April 2004.

Government of East Timor (2004). *Our national vision.* http://www.gov-rdtl.org. Accessed on 6th April 2004.

Hajek J (2000). 'Language planning and the sociolinguistic environment in East Timor: colonial practice and changing language ecologies.' *Current Issues in Language Planning 1(3),* 400–414.

Hezel F X (1984). 'Schools in Micronesia prior to American administration.' *Pacific Studies 8(1),* 95–111. Retrieved November 30, 2004 from. http://www.micsem.org.

Hindson C (1985). 'Post-primary school non-academic alternatives: a south Pacific study.' *Comparative Education 21(2),* 135–156.

Hull G (2000). *East Timor: identity, language and education policy.* East Timor: Address to the CNRT National Congress in Dili, East Timor, 25 August 2000.

Jourdan C (1990). 'Solomon's Pijin: an unrecognized national language.' In Baldauf R B Jr & Luke A (eds.) *Language planning and education in Australasia and the South Pacific.* Clevedon: Multilingual Matters. 166–181.

Kaplan R B (1993). 'Conquest of paradise: language planning in New Zealand.' In Hoey M & Fox G (eds.) *Data, description, discourse: festschrift for John Sinclair on his 60th birthday.* London: HarperCollins. 151–172.

Kaplan R B & Baldauf R B Jr (2003). 'Language planning in Australia.' In *Language and language-in-education planning in the Pacific Basin.* Dordrecht: Kluwer. 143–166.

Kaplan R B & Baldauf R B Jr (2003). 'Language planning in Indonesia.' In *Language and language-in-education planning in the Pacific Basin.* Dordrecht: Kluwer. 83–102.

Kaplan R B & Baldauf R B Jr (2003). 'Language planning in Melanesia.' In *Language and language-in-education planning in the Pacific Basin.* Dordrecht: Kluwer. 185–200.

Kaplan R B & Baldauf R B Jr (2003). 'Language planning in New Zealand.' In *Language and language-in-education planning in the Pacific Basin.* Dordrecht: Kluwer. 167–183.

Kennedy T (1984). 'The Cook Islands.' In Thomas R M & Postlethwaite T N (eds.) *Schooling in the Pacific Islands: colonies in transition.* Oxford: Pergamon. 262–294.

Laurens C & Vareille J (1984). 'New Caledonia and the Society Islands.' In Thomas R M & Postlethwaite T N (eds.) *Schooling in the Pacific Islands: colonies in transition.* Oxford: Pergamon. 109–142.

Lotherington H (1998). 'Trends and tensions in post-colonial language education in the South Pacific.' *International Journal of Bilingual Education and Bilingualism 1(1), 65–75.*

Lynch J & Mugler F (2002). Pacific languages of the University of the South Pacific. *Current Issues in Language Planning 3(1),* 76–81.

McCarty T L (2001). 'Between possibility and constraint: indigenous language education, planning and policy in the United States.' In Tollefson J W (ed.) *Language policies in education: critical issues.* Mahwah, NJ: Lawrence Erlbaum. 285–307.

Mangubhai F (2002). 'Language-in-education policies in the South Pacific: some possibilities for consideration.' *Journal of Multilingual and Multicultural Development 23(6),* 490–511.

Mangubhai F & Mugler F (2003). 'The language situation in Fiji.' *Current Issues in Language Planning 4,* 185–280.

Mugler F (1996). '"Vernacular" language teaching in Fiji.' In Mugler F & Lynch J (eds.) *Pacific languages in education.* Suva: Institute of Pacific Studies, University of the Pacific. 273–287.

Mugler F & Lynch J (eds.) (1996). *Pacific languages in education.* Suva: Institute of Pacific Studies, University of the Pacific.

Mühlhäusler P (1998). 'How creoloid can you get?' *Journal of Pidgin and Creole Languages 13(2),* 355–371.

National Office of Overseas Skills Recognition (NOOSR) (1995). *Country education profiles: Pacific Islands I: Solomon Islands, Tonga – a comparative study.* Canberra: A. G. P. S.

National Office of Overseas Skills Recognition (NOOSR) (1995). *Country education profiles: Pacific Islands II: Cook Islands, Kiribati, Western Samoa – a comparative study.* Canberra: A. G. P. S.

NOOSER (1997). *Country education profiles: Fiji – a comparative study* (3rd edn.). Canberra: A. G. P. S.

Norick J S (1981). *A Tuvalu dictionary, Tuvalu–English, English–Tuvalu.* New Haven, CT: Human Relation Area Files Book.

Pearson M (2004). *Kiribati: education.* http://www.wysiwyg.co.nz. Originally by Jonathan Willis-Richards.

Peddie R (1997). 'Why are we waiting? Languages policy development in New Zealand.' In Eggington W G & Wren H (eds.) *Language policy: dominant English, pluralist challenges.* Amsterdam: John Benjamins. 121–146.

Pine P & Savage W (1989). 'Marshallese and English: evidence for an immersion model of education in the

Republic of the Marshall Islands.' *World Englishes 8(1),* 83–94.
Renandya W A (2003). 'Indonesia.' In Ho W K & Wong R Y L (eds.) *Language policies and language education in East Asia.* Singapore: Singapore Association for Applied Linguistics.
Rensch K H (1990). 'The delayed impact: postcolonial language problems in the French Overseas Territory Wallis and Futuna (Central Polynesia).' *Language Problems and Language Planning 14(3),* 224–236.
Shameem N (2002). 'Multilingual proficiency in Fiji primary schools.' *Journal of Multilingual and Multicultural Development 23(5),* 388–407.
Siegel J (1987). 'Language at the University of the South Pacific.' *Indian Journal of Applied Linguistics 13(1),* 53–63.
Siegel J (1997). 'Formal vs. non-formal vernacular education: the education reform in Papua New Guinea.' *Journal of Multilingual and Multicultural Development 18(3),* 206–222.
Smolicz J J & Secombe M J (2003). 'Assimilation or pluralism? Changing policies for minority languages education in Australia.' *Language Policy 2(1),* 3–25.
Spencer M (1996). 'And what of the languages of Micronesia?' In Mugler F & Lynch J (eds.) *Pacific languages in education.* Suva: Institute of Pacific Studies, University of the Pacific. 10–35.
Spolsky B, Engelbrecht G & Ortiz L (1983). 'Religious, political and educational factors in the development of biliteracy in the Kingdom of Tonga.' *Journal of Multilingual and Multicultural Development 4(6),* 459–469.
Thomas R M (1984a). Tonga. In Thomas R M & Postlethwaite T N (eds.) *Schooling in the Pacific Islands: colonies in transition.* Oxford: Pergamon. 236–261.
Thomas R M (1984b). 'The U.S. Trust Territory of the Pacific islands (Micronesia).' In Thomas R M & Postlethwaite T N (eds.) *Schooling in the Pacific Islands: colonies in transition.* Oxford: Pergamon. 67–107.
Trann Ngoc A (1996). 'Situation et appui a l'enseignement des langues maternelles au Vanuatu, a Wallis et Futuna et en Nouvelle-Caledonie dans le Pacifique-Sud francophone' [The situation and support for native-language instruction in Vanuatu, Wallis and Futuna, and New Caledonia in French-speaking South Pacific]. *Etudes de Linguistique Appliquee 103 (July–Sept),* 323–328.
UNESCO (2004). www.unesco.org. Accessed on 26th March 2004.
Underwood R A (1989a). 'Education and Chamorro identity on Guam.' *Ethnies: Droits de l'homme et peuples autochtones 8/9/10,* 36–40.
Underwood R A (1989b). 'English and Chamorro on Guam.' *World Englishes 8(1),* 73–82.
Vetter R (1991). 'Discourses across literacies: personal letter writing in a Tuvaluan context.' *Language and Education 5,* 125–145.
Wikipedia (2004). http://en.wikipedia.org. Accessed on 28th March 2004.
World Bank (2003). *World development indicator 2003.* http://devdata.worldbank.org. Accessed 7 April 2004.
Wurm S A (1999). 'Language revivalism and revitalization in Pacific and Asian Areas.' *International Journal of the Sociology of Language 137,* 163–172.
Yunick S Jr (2000). 'Linguistics, TESL, and language planning in Micronesia.' *Studies in the Linguistic Sciences 30(1),* 183–200.

Language Education Policy in China

M Zhou, Dickinson College, Carlisle, PA, USA

Introduction

Language education policy in the People's Republic of China (PRC) consists of the PRC Constitution provisions and education-oriented legislation, the PRC government's executive regulations and decrees, and the Chinese Communist Party's (CCP) policies and directives. The relationship among the three is rather complicated, but as a rule the CCP's policies and directives have been underpinning the other two and, sometimes, above them, at least as far as language education policy is concerned (see Blachford, 2004; Zhou, 2004a).

As a multinational and multilingual state (Mackerras, 1994; Ramsey, 1987), the PRC has been struggling between accommodation and integration, a struggle that has been conditioned by the CCP's ideology, the Cold War, and globalization (Dreyer, 1976; Zhou, 2001, 2003: 36–98). This has resulted in pluralistic (1949–1958), integrationist (1958–1977), and bilingual/multilingual (1978–present) stages of language education policy development. Regardless of the changes, China always faces challenges as how to implement its national policies and how to make adjustments to local conditions (Lin, 1997; Xiao, 1998; Zhou, 2004b).

The Pluralistic Stage, 1949–1957

The pluralistic stage was characterized by accommodationism, which allowed the development of independent language education policies for Chinese and minority languages.

Republic of the Marshall Islands.' *World Englishes 8(1),* 83–94.
Renandya W A (2003). 'Indonesia.' In Ho W K & Wong R Y L (eds.) *Language policies and language education in East Asia.* Singapore: Singapore Association for Applied Linguistics.
Rensch K H (1990). 'The delayed impact: postcolonial language problems in the French Overseas Territory Wallis and Futuna (Central Polynesia).' *Language Problems and Language Planning 14(3),* 224–236.
Shameem N (2002). 'Multilingual proficiency in Fiji primary schools.' *Journal of Multilingual and Multicultural Development 23(5),* 388–407.
Siegel J (1987). 'Language at the University of the South Pacific.' *Indian Journal of Applied Linguistics 13(1),* 53–63.
Siegel J (1997). 'Formal vs. non-formal vernacular education: the education reform in Papua New Guinea.' *Journal of Multilingual and Multicultural Development 18(3),* 206–222.
Smolicz J J & Secombe M J (2003). 'Assimilation or pluralism? Changing policies for minority languages education in Australia.' *Language Policy 2(1),* 3–25.
Spencer M (1996). 'And what of the languages of Micronesia?' In Mugler F & Lynch J (eds.) *Pacific languages in education.* Suva: Institute of Pacific Studies, University of the Pacific. 10–35.
Spolsky B, Engelbrecht G & Ortiz L (1983). 'Religious, political and educational factors in the development of biliteracy in the Kingdom of Tonga.' *Journal of Multilingual and Multicultural Development 4(6),* 459–469.
Thomas R M (1984a). Tonga. In Thomas R M & Postlethwaite T N (eds.) *Schooling in the Pacific Islands: colonies in transition.* Oxford: Pergamon. 236–261.
Thomas R M (1984b). 'The U.S. Trust Territory of the Pacific islands (Micronesia).' In Thomas R M & Postlethwaite T N (eds.) *Schooling in the Pacific Islands: colonies in transition.* Oxford: Pergamon. 67–107.
Trann Ngoc A (1996). 'Situation et appui a l'enseignement des langues maternelles au Vanuatu, a Wallis et Futuna et en Nouvelle-Caledonie dans le Pacifique-Sud francophone' [The situation and support for native-language instruction in Vanuatu, Wallis and Futuna, and New Caledonia in French-speaking South Pacific]. *Etudes de Linguistique Appliquee 103 (July–Sept),* 323–328.
UNESCO (2004). www.unesco.org. Accessed on 26th March 2004.
Underwood R A (1989a). 'Education and Chamorro identity on Guam.' *Ethnies: Droits de l'homme et peuples autochtones 8/9/10,* 36–40.
Underwood R A (1989b). 'English and Chamorro on Guam.' *World Englishes 8(1),* 73–82.
Vetter R (1991). 'Discourses across literacies: personal letter writing in a Tuvaluan context.' *Language and Education 5,* 125–145.
Wikipedia (2004). http://en.wikipedia.org. Accessed on 28th March 2004.
World Bank (2003). *World development indicator 2003.* http://devdata.worldbank.org. Accessed 7 April 2004.
Wurm S A (1999). 'Language revivalism and revitalization in Pacific and Asian Areas.' *International Journal of the Sociology of Language 137,* 163–172.
Yunick S Jr (2000). 'Linguistics, TESL, and language planning in Micronesia.' *Studies in the Linguistic Sciences 30(1),* 183–200.

Language Education Policy in China

M Zhou, Dickinson College, Carlisle, PA, USA

Introduction

Language education policy in the People's Republic of China (PRC) consists of the PRC Constitution provisions and education-oriented legislation, the PRC government's executive regulations and decrees, and the Chinese Communist Party's (CCP) policies and directives. The relationship among the three is rather complicated, but as a rule the CCP's policies and directives have been underpinning the other two and, sometimes, above them, at least as far as language education policy is concerned (see Blachford, 2004; Zhou, 2004a).

As a multinational and multilingual state (Mackerras, 1994; Ramsey, 1987), the PRC has been struggling between accommodation and integration, a struggle that has been conditioned by the CCP's ideology, the Cold War, and globalization (Dreyer, 1976; Zhou, 2001, 2003: 36–98). This has resulted in pluralistic (1949–1958), integrationist (1958–1977), and bilingual/multilingual (1978–present) stages of language education policy development. Regardless of the changes, China always faces challenges as how to implement its national policies and how to make adjustments to local conditions (Lin, 1997; Xiao, 1998; Zhou, 2004b).

The Pluralistic Stage, 1949–1957

The pluralistic stage was characterized by accommodationism, which allowed the development of independent language education policies for Chinese and minority languages.

From the beginning, the PRC provided adequate constitutional protection of linguistic pluralism in education (Zhou, 2004a). Passed as the provisional constitution in September 1949, the Common Program (Article 53) guaranteed minority nationalities freedom to use and develop their languages and scripts (Chen, 1967: 34–45). In 1954 the Chinese National People's Congress passed the first PRC Constitution, Article 3 of which guaranteed the same freedom (Chen, 1967: 75–92).

To implement the Constitutional guarantees, the Ministry of Education made a three-point decision on minority language education in 1951 (China, 1991: 28–40):

1. In minority communities with commonly used languages, such as Mongol, Korean, Tibetan, Uygur, Kazak, etc., primary and secondary schools should use native languages as the medium of instruction for all courses;
2. Minority communities without writing systems or with imperfect ones should begin to create or reform their writing systems, and at the same time adopt Chinese or a customarily used language as the medium of instruction on a voluntary basis; and
3. Chinese should be offered only on demand and on a voluntary basis in schools in minority communities.

The national language, known as *Putonghua* (Standard Modern Chinese), became a top priority when the first national conference on script reform recommended its promotion in October 1955. In November the Ministry of Education disseminated a decree (1996: 48–55):

1. Primary, secondary, and normal schools throughout the country should gradually use Putonghua as the medium of instruction and teach students how to speak it, beginning from language classrooms and spreading to all subject classrooms; and
2. Putonghua should be used in Chinese language classes in schools in minority communities.

It was firmly supported by a CCP directive in January 1956, and in February followed by the State Council's decree that specified (China, 1996: 7–15):

1. Starting in winter 1956 Putonghua must be taught in all primary and secondary schools (excluding minority schools) throughout the country.
2. By 1960 (a) third graders and above and secondary and normal school students should be basically able to speak Putonghua; (b) primary, secondary, and normal school teachers should be able to use Putonghua as the medium of instruction; and (c) graduating secondary school and college students and young teachers should undergo extensive training if they could not speak Putonghua.
3. All military schools should use Putonghua as the medium of instruction, and so should all evening and literacy schools.
4. The Ministry of Education should hold frequent training classes for teachers and educational administrators.
5. Putonghua should be promoted among the Han people (the majority) residing in minority communities and be taught in Chinese classes in minority schools.

This decree was the basis for Chinese education policy in the 1950s and for years to come. Meanwhile, the ministry also instructed schools to use simplified Chinese characters in classrooms, textbooks, and school documents, because traditional ones were considered a hindrance to rapid literacy development (China, 1996: 56; Rohsenow, 2004).

By 1956 education policies for Chinese and minority languages still remained rather independent. But in 1957 they were about to merge as Putonghua promotion gathered steam and *Pinyin,* a Roman alphabet intended to assist Chinese script pronunciation and ultimately to replace it and with a legacy long before the PRC, was published for public discussion before its formal promulgation in 1958 (Chen, 1999: 186–189; DeFrancis, 1950; Norman, 1988: 263; Rohsenow, 2004).

The Integrationist Stage, 1958–1976

Undergoing the Great Leap Forward (1958–1960) and the Cultural Revolution (1966–1976) (see Spence, 1991:574–583 and 602–652), China at the integrationist stage aggressively promoted Chinese and frequently suppressed minority languages in minority schools, although the 1954 Constitution had always been in effect and the simplified 1975 Constitution still protected minority language rights.

When Pinyin was officially promulgated at the beginning of 1958 the Chinese government proposed that it serve as the common base for creating and reforming writing systems for minority languages, in order to facilitate cultural exchange and communication among nationalities (Seybolt and Chiang, 1979: 228–243; Zhou, 2003: 203–208). Moreover, in March 1958 the Ministry of Education requested that all primary, secondary, and normal schools, including minority ones, start teaching Pinyin in fall 1958.

At the same time, the State Nationalities Affairs Commission made six recommendations to hasten integration, four of which were directly related to language education:

1. Chinese (Putonghua) should be offered in minority primary schools (allowing the provincial and

autonomous governments to decide the starting grade).
2. Han students should also learn minority languages when studying in schools that used minority languages as the medium of instruction.
3. Chinese or other languages might be used in literacy campaigns and schools in minority communities that had not created their writing systems yet or did not need such systems.
4. Chinese or other minority scripts might be used in literacy campaigns on a voluntary basis even if a minority group had its own script.

This policy was almost consistently implemented during this period, leading to bi/multilingualism in many minority schools and communities, whereas minority language education was only occasionally encouraged. Official Putonghua promotion in schools saw some highs in the form of nationwide model demonstration conferences at which model teachers and learners exchanged their successful experiences with conference participants, but the promotion was interrupted during the Cultural Revolution.

The Bi/multilingual Stage, 1977–Present

This stage has strengthened language education legislation to consolidate the new linguistic order with Chinese as the national language and minority languages as local ones, after the revival of the PRC's early pluralistic policies in the late 1970s and early 1980s.

After the Cultural Revolution China's linguistic landscape was transformed to bi/multilingualism (Chinese and minority languages) in many minority communities and bi/multidialectalism (Putonghua, its variants, and Chinese dialects) in almost all Chinese-speaking communities (Zhou, 2003; Saillard, 2004). Language education policy had to accommodate this diversity.

In 1980, with authorization from the State Council, the Ministry of Education and State Nationalities Affairs Commission instructed that minorities could use their own languages and scripts as the medium of instruction in schools and they should master both their native languages and Chinese (China, 1991: 91–96). Two years later, in 1982, Putonghua's position as the national commonly used language was adopted in the newly amended PRC Constitution (Article 19) (China, 1999), leading to a contradiction between the constitutional freedom to use minority languages and the constitutional obligation to use the national language, a contradiction that has since been reflected in national (and local) laws on education.

The PRC autonomy law for minority communities, which was passed in 1984, states that (a) autonomous governments can specify the medium of instruction for local schools in accordance with central laws and policies (Article 36) and (b) when conditions permit schools with a predominantly minority student body should use minority languages both in textbooks and instruction but still promote national commonly used Putonghua (Article 37) (China, 1998a). Passed in 1986, the PRC compulsory education law stipulates that schools should promote national commonly used Putonghua, and minority schools may adopt commonly used minority languages and scripts as the medium of instruction (Sun and Gao, 1996: 91–93). The PRC Education Law (passed in 1995) further specifies Putonghua as the primary medium of instruction (Article 12), with a similar clause on minority language use (China, 1998b: 2):

> Chinese language and script is the primary medium of instruction for schools and other educational organizations, but those with a student body of predominantly minority origin may use native or local commonly used languages and scripts as a medium of instruction.

All schools and other educational organizations should promote national commonly used Putonghua and standardized Chinese characters in their teaching. These laws delimit the constitutional freedom of minority language use by adding 'when conditions permit' and 'commonly used.'

To use Putonghua as the primary medium of instruction meets challenges not only in minority schools but also in Chinese schools where varieties of Putonghua and Chinese dialects are spoken. To accommodate the varieties of Putonghua in Chinese and minority schools, in 1986 the State Language Commission proposed to recognize three levels of Putonghua, an unprecedented recognition of different levels of a standard language in the history of status planning (China, 1987: 25–26; Saillard, 2004). In 1994 the Ministry of Education and State Language Commission published a regulation on how to measure Putonghua levels (three levels with A and B grades in each; 1A = highest; 3B = lowest) and who should attain which level (China, 1998b: 2072–2076). Schools, teachers, and students are all regularly assessed for their Putonghua proficiency individually and collectively (China, 1998b: 396–398). From 1995 on, primary and secondary school teachers must gradually reach Levels 1 or 2 before they are assigned teaching duties. The 2001 language law (passed in 2000 and effective in 2001) has confirmed Putonghua as the primary language of instruction (Article 10), minority languages as a legally permissible medium of instruction (Articles 8 and

10), and levels of Putonghua and level requirements for teachers, legislating a long existing linguistic order in schools and in the society (China, 2001: 4–9; Rohsenow, 2004).

Conclusion

Since 1949 the PRC has provided adequate constitutional and legislative protection of linguistic diversity in education with Putonghua as the primary medium of instruction and minority languages as local ones. Minority languages are used as a (supplementary) medium of instruction from primary schools to colleges, whereas Chinese is offered as subject courses and as a medium of instruction from primary schools to colleges in minority communities.

However, the constitutional provisions meet two serious challenges in the implementation. First, the language education policy was hijacked by the CCP's political agenda from the late 1950s to the mid–1970s, and has often been delimited by central and local executive decrees and regulations (Zhou, 2004a). Second, the policy has always faced the problem of how to adjust to and successfully accommodate local multilingual and multicultural conditions (Bradley and Bradley, 2002; Hansen, 1999; Harrell, 1993; Postiglione, 1999; Zhou, 2004b).

See also: Language Policy in Multilingual Educational Contexts; Minority Languages: Education.

Bibliography

Blachford D R (2004). 'Language spread versus language maintenance: Policy making and implementation process.' In Zhou M (ed.). 99–122.

Bradley D & Bradley M (2002). 'Language policy and language maintenance: Yi in China.' In Bradley D & Bradley M (eds.) *Language endangerment and language maintenance*. London: Routledge Curzon. 77–97.

Chen P (1999). *Modern Chinese: history and sociolinguistics*. Cambridge: Cambridge University Press.

Chen T H E (1967). *The Chinese communist regime: documents and commentary.* New York: Praeger.

China (1958). 'Wei zhengque guanche zhixing dang de minzu yuwen gongzuo fangzhen er fendou [Strive to correctly implement the Party's minority language and script policy].' *Minzu Yanjiu 3*[Nationalities Research]: 1–9.

China (1987). *Xinshiqi de yuyan wenzi gongzuo – quanguo yuyan wenzi gongzuo huiyi wenjian huibian* [Language and script work in the new era: collection of documents from the national conference on language and script work]. Beijing: Language Press.

China (1991). *Shaoshu minzu jiaoyu gongzuo wenjian xuanbian 1949–1988* [Selected documents of the Central Government on minority education: 1949–1988]. Hohhot: Inner Mongolian Education Press.

China (1996). *Guojia yuyan wenzi zhengce fagui huibian* (1949–1995) [Collection of state language and script policies and regulations: 1949–1995]. Beijing: Language Press.

China (1998a). *Zhonghua Renmin Gongheguo Minzu Quyu Zizhifa* [The PRC Law of Regional Autonomy for Minority Nationalities]. Beijing: Chinese Law Press.

China (1998b). *Zhonghua Renmin Gongheguo xiangxing jiaoyu fagui huibian 1990–1995* [Collection of current education laws and regulations of the PRC, 1990–1995]. Beijing: People's Education Press.

China (1999). *Zhonghua Renmin Gongheguo Xianfa* [The Constitution of the People's Republic of China]. Beijing: People's Press.

China (2001). *Zhonghua renmin gongheguo guojia tongyong yuyan wenzi fa xuexi duben* [A reader of the PRC law on the national commonly-used language and script]. Beijing: Yuyan Press.

DeFrancis J (1950). *Nationalism and language reform in China*. Princeton, NJ: Princeton University Press.

Dreyer J T (1976). *China's forty million.* Cambridge, MA: Harvard University Press.

Hansen M H (1999). *Lessons in being Chinese: minority education and ethnic identity in Southwest China.* Seattle: University of Washington Press.

Harrell Stevan (1993). 'Linguistics and hegemony in China.' *International Journal of the Sociology of Language 103,* 97–114.

Lin J (1997). Policies and practices of bilingual education for minorities in China. *Journal of Multilingual and Multicultural Development 18,* 193–205.

Mackerras C (1994). *China's minorities: integration and modernization in the twentieth century.* Hong Kong: Oxford University Press.

Norman J (1988). *Chinese.* Cambridge: Cambridge University Press.

Postiglione G A (ed.) (1999). *China's national minority education: culture, schooling, and development.* New York: Falmer Press.

Ramsey S R (1987). *The languages of China*. Princeton, NJ: Princeton University Press.

Rohsenow J (2004). 'Fifty years of script and written Language reform in the PRC: the genesis of the Language Law of 2001.' In Zhou M (ed.). 21–41.

Saillard C (2004). 'On the promotion of Putonghua in China: how a standard language becomes a vernacular.' In Zhou M (ed.). 163–176.

Seybolt P J & Chiang G K (eds.) (1979). *Language reform in China: documents and commentary.* Armonk, NY: M. E. Sharpe.

Spence J D (1990). *The search for modern China.* New York: W. W. Norton.

Sun W Z & Gao Z X (eds.) (1996). *Zhonghua Renmin Gongheguo jiaoyufa shouce* [A handbook of education laws of the People's Republic of China]. Beijing: Chinese Statistical Publishing House.

Xiao H (1998). 'Minority languages in Dehong, China: policy and reality.' *Journal of Multilingual and Multicultural Development 19,* 221–235.

Zhou M (2001). 'The politics of bilingual education in the People's Republic of China since 1949.' *Bilingual Research Journal 25(1/2),* 147–171.
Zhou M (2003). *Multilingualism in China: the politics of writing reforms for minority languages 1949–2002.* Berlin/New York: Mouton de Gruyter.
Zhou M (2004a). 'Minority language policy in China: equality in theory and inequality in practice.' In Zhou M (ed.). 73–96.
Zhou M (ed.) (2004b). *Language policy in the People's Republic of China: theory and practice since 1949.* Norwell, MA: Kluwer.

Language Education Policy in Europe

J Trim, Cambridge, UK

Language policy in Europe, as elsewhere, is a highly complex matter. At one end, we have the attempts of those in authority to impose patterns of behavior on those subject to that authority. Some (e.g., Ager, 2001) would confine the term *policy* to the language behavior of those in power. However, the actual use and learning of languages by individuals everywhere is subject to the influence of many interacting agents: parents, family and friends, peer groups, teachers, information and entertainment media, authors both living and dead, etc. Ultimately, it is the decisions, conscious and unconscious, made by countless individuals in response to these influences that provide the reality of language policy. Inevitably, almost all those decisions are unobserved and perhaps unobservable. We concentrate attention on decisions at a high level, affecting large numbers of individuals, but we should not lose sight of the fact that it is only to the extent that individuals act in accordance with such decisions as interpreted and relayed by intermediates that policies are effective.

In the permissive atmosphere of the pluralist democracies that characterize contemporary Europe, political decision-making is increasingly decentralized. This is somewhat paradoxical, because at the same time, society is becoming more homogeneous and more tightly knit. Individuals find themselves participating in denser and more widely cast networks of communication than ever before.

The Historical Background

The basic language pattern of Europe today is the result of the large-scale migration movements following the collapse of the Roman Empire in the 5th century A.D. and the settlement of a large part of the European Continent by speakers of Germanic and Slav dialects. During the centuries that followed, poor communication led to a broad spectrum of largely mutually unintelligible forms of speech, tolerated by the thin administrative superstratum in the loose, shifting, dynastic feudal empires of the time. By the late medieval period, communications had much improved. The emergence of urban centers and the strengthening of social, economic, and political organization, with the rise of the middle class, started a process of crystallization, accelerated by the invention of printing, the growth of mother tongue literacy, and mass education. The modern nation state, epitomized by France, has seen itself as monolingual and monocultural. *Egalité* and *fraternité* allowed no place for minority languages and cultures. Compulsory education brought mass literacy in a standardized written language, based generally on the usage of an administrative class in a capital city. In the course of the 19th and 20th centuries, the older multilingual empires disappeared, and the patchwork of independent nation states became the norm. Politically fragmented language communities unified (Italy, Germany). Within multilingual states, local dialects and minority languages came under heavy pressure and were threatened with extinction. In self-defense, language-based separatist and irredentist movements became strong.

Current Development

However, at the same time as this historical development has been reaching its culmination, further dramatic improvements in communication have made it obsolescent. While 19th century technical developments (railway travel, postal services, mass circulation books and newspapers) reinforced the internal unity of the nation state within enforceable frontiers, developments of the 20th century (air transport, continental motorway networks, electronic media) have transcended those frontiers. Industry, commerce, finance, and the expanding leisure industries are coming to be organized on an international, even global, scale. Personal mobility for vocational, educational, and private purposes now entails travel across political and linguistic boundaries on a large scale. As a result, the language policy issues confronting national authorities in Europe arise from their

Zhou M (2001). 'The politics of bilingual education in the People's Republic of China since 1949.' *Bilingual Research Journal 25(1/2),* 147–171.
Zhou M (2003). *Multilingualism in China: the politics of writing reforms for minority languages 1949–2002.* Berlin/New York: Mouton de Gruyter.
Zhou M (2004a). 'Minority language policy in China: equality in theory and inequality in practice.' In Zhou M (ed.). 73–96.
Zhou M (ed.) (2004b). *Language policy in the People's Republic of China: theory and practice since 1949.* Norwell, MA: Kluwer.

Language Education Policy in Europe

J Trim, Cambridge, UK

Language policy in Europe, as elsewhere, is a highly complex matter. At one end, we have the attempts of those in authority to impose patterns of behavior on those subject to that authority. Some (e.g., Ager, 2001) would confine the term *policy* to the language behavior of those in power. However, the actual use and learning of languages by individuals everywhere is subject to the influence of many interacting agents: parents, family and friends, peer groups, teachers, information and entertainment media, authors both living and dead, etc. Ultimately, it is the decisions, conscious and unconscious, made by countless individuals in response to these influences that provide the reality of language policy. Inevitably, almost all those decisions are unobserved and perhaps unobservable. We concentrate attention on decisions at a high level, affecting large numbers of individuals, but we should not lose sight of the fact that it is only to the extent that individuals act in accordance with such decisions as interpreted and relayed by intermediates that policies are effective.

In the permissive atmosphere of the pluralist democracies that characterize contemporary Europe, political decision-making is increasingly decentralized. This is somewhat paradoxical, because at the same time, society is becoming more homogeneous and more tightly knit. Individuals find themselves participating in denser and more widely cast networks of communication than ever before.

The Historical Background

The basic language pattern of Europe today is the result of the large-scale migration movements following the collapse of the Roman Empire in the 5th century A.D. and the settlement of a large part of the European Continent by speakers of Germanic and Slav dialects. During the centuries that followed, poor communication led to a broad spectrum of largely mutually unintelligible forms of speech, tolerated by the thin administrative superstratum in the loose, shifting, dynastic feudal empires of the time. By the late medieval period, communications had much improved. The emergence of urban centers and the strengthening of social, economic, and political organization, with the rise of the middle class, started a process of crystallization, accelerated by the invention of printing, the growth of mother tongue literacy, and mass education. The modern nation state, epitomized by France, has seen itself as monolingual and monocultural. *Egalité* and *fraternité* allowed no place for minority languages and cultures. Compulsory education brought mass literacy in a standardized written language, based generally on the usage of an administrative class in a capital city. In the course of the 19th and 20th centuries, the older multilingual empires disappeared, and the patchwork of independent nation states became the norm. Politically fragmented language communities unified (Italy, Germany). Within multilingual states, local dialects and minority languages came under heavy pressure and were threatened with extinction. In self-defense, language-based separatist and irredentist movements became strong.

Current Development

However, at the same time as this historical development has been reaching its culmination, further dramatic improvements in communication have made it obsolescent. While 19th century technical developments (railway travel, postal services, mass circulation books and newspapers) reinforced the internal unity of the nation state within enforceable frontiers, developments of the 20th century (air transport, continental motorway networks, electronic media) have transcended those frontiers. Industry, commerce, finance, and the expanding leisure industries are coming to be organized on an international, even global, scale. Personal mobility for vocational, educational, and private purposes now entails travel across political and linguistic boundaries on a large scale. As a result, the language policy issues confronting national authorities in Europe arise from their

being caught as a 'piggy in the middle' between the demand from minority communities for greater autonomy on the one hand, and the pressures of globalization on the other.

In all, up to some 200 indigenous languages are currently used in Europe, though the number is indefinitely higher if the languages of recent immigrants are taken into account. They vary greatly in the number of native speakers, speakers of the language as a second or as a foreign language, and in their role and status in society. It may be useful to follow the hierarchical distinction proposed by Abram de Swaan (2001) between peripheral, central, supercentral, and hypercentral languages.

Peripheral languages, to which most of the world's languages belong, generally occupy a small geographical area, have relatively few native speakers, and are learned as a second language only by outsiders marrying into or otherwise closely connected with the community concerned. They have virtually no role in organized foreign language teaching. They are used mainly for interpersonal spoken communication and often have no standard written form. They tend not to be used in public administration, in broadcast media, or in the education system. Their speakers are more or less obliged to be bilingual, especially when they have minority status within a larger political entity. Often, bilingualism leads to diglossia, with the peripheral language increasingly confined to domestic life and community functions, often cultural or religious. Under modern conditions, they are often at the limit of viability, used more by older people than by the younger generation, and are in danger of extinction. A surprising number of peripheral languages are in this position in Europe, largely unknown to the outside world.

Central languages are generally the official or *de facto* state languages of nations large and small, each existing in a standard written form used in newspapers and normal publications, in broadcast media, in public administration, and as the medium of instruction throughout the public education system. There is usually a standard spoken form based on the practice of the educated urban middle class, especially in the capital city. There are strong pressures upon all long-term residents, perhaps a legal requirement upon applicants for citizenship, to acquire and use it as a second language. Central languages have a place in foreign language teaching, but it is often marginal.

Supercentral languages are those central languages that are also the medium of international communication for a group of central languages. They are few in number, but they have large numbers of speakers, both native and non-native. In Europe, English, French, German, Spanish, and Russian have this status, though that of Spanish is mainly extra-European. They have an established position in foreign language teaching within their respective spheres of influence.

A hypercentral language acts as the medium of international communication, not for one restricted group of central languages, but for all. Its currency is global and an aspect of globalization, which both promotes and is reinforced by it. At present, English appears to be firmly established as the sole language in this role, in Europe, as elsewhere.

Minority Languages

The languages of Europe fall fairly well into one of those four categories, though each has its special characteristics distinguishing it from the stereotypical character sketched above, especially in the case of the peripheral languages. For instance, during the rise of the nation state, the use of minority languages was discouraged, even forbidden, and the speakers oppressed, especially in the educational system. Notoriously, clappers were attached to the tongues of children caught speaking Welsh at school. In France, the ideology of the Revolution maintained that standard French was the sole national language required by the values of equality and fraternity, which excluded the very concept of minorities. This model was followed by other emerging national states such as Turkey, where the existence of the Kurdish language was officially denied for decades. Following the Spanish Civil War, the Franco regime forbade the use of Catalan, Basque, and Galician in Northern Spain as divisive. At the same time, large-scale internal migration was encouraged, so that the number of children in Catalonia entering primary school with Catalan as their home language fell to 25%. All three languages now enjoy the status of official languages alongside Spanish in the autonomous regions concerned.

However, though many minority languages were pushed to the limits of viability, few actually succumbed. Most, if not all, European states still contain significant minority language populations. Attitudes have changed dramatically. Territorial minority heritage languages are now not simply tolerated, but positively encouraged. In a number of cases, small states have achieved or retained independence, raising a less widely used language to the status of a state language, retaining some characteristics of a peripheral language but others of a central language. The mix varies. Letzeburgisch is the everyday spoken language of the indigenous population of Luxembourg. It is dialectally differentiated, and despite a brief attempt at the end of the World War II to use a

standardized form for official purposes, it has no established standard written form. French and German are used as a matter of state policy for educational and administrative purposes, primary education being conducted in German and secondary education in French. In addition, no less than 36.7% of the resident population are economic migrants from other European countries, as are 9% of residents in Germany, Austria, and Belgium, as well as 6.5% in Sweden. There are appreciable numbers of migrants from other continents in the ex-imperial countries of France (6.2%), the United Kingdom (3.5%), and The Netherlands (4.2%).

The status of minority languages and their role in public life, administration, and education varies widely. The *Universal declaration of language rights* (1998), recognized by UNESCO, makes the most far-reaching claims for the use of mother tongue throughout the educational system, in courts, and in other contacts with officialdom at all administrative levels. The European Bureau for Lesser Used Languages (EBLUL) is supported by the European Union (EU) and is mostly concerned with the protection and promotion of territorial minorities. The Council of Europe has produced the *European charter for regional and minority languages* (1992). Its provisions are mostly qualified and non-mandatory, but even so it has been signed by 28 of the 46 member states (though ratified by only 15). Those signatories do not include countries where the problems are most acute, such as the Baltic States, especially Latvia where Russification had reduced the national language to minority status, especially in major cities. While Russian continues to be widely used in daily life and in much of the education system, citizenship and employment laws have been enforced that aim to restore Latvian as the normal language in public life.

In 2001, a native speaker of Arumanian (Vlach) was found guilty by a Greek court for spreading "false information on minority languages in Greece," and of causing disorder by claiming that a number of minority languages (Arumanian, Pomek, Macedonian, Albanian, and Turkish) are spoken in Greece. The verdict was, however, overturned on appeal.

Some countries (Ireland, Finland, Switzerland, Malta, Belgium) are officially bilingual. In others (e.g., United Kingdom, Italy, Spain), bilingualism is officially established in autonomous ethnic minority regions. Again, the actual relations vary widely. In Ireland, the everyday use of Irish is restricted to a dwindling Gaeltacht area in the West. Elsewhere in Ireland, while public signs are bilingual and parents strongly support the teaching of Irish in schools, they do not wish to speak it in daily life, which is as monolingually English as it is in England. In Belgium, where historically French was displacing Dutch in urban centers, reducing it to a rural patois, Flemish nationalism has established separate language hegemonies for each region (including a small German-speaking area around Eupen). Bilingual education, in the sense of the use of another language as the medium of instruction, is illegal.

In Finland, the role of Swedish is declining, with its use confined to shrinking coastal communities and its role in education giving way to English. In Switzerland, life in the three main communities is conducted monolingually, with Rhaeto-romanisch a protected minority language within the German-speaking area. The decision by Zürich to give the teaching of English priority over that of French caused outrage. In Malta and to some extent in Cyprus, the survival of English, the ex-colonial language, has been seen as an economic benefit. In officially bilingual regions, mother-tongue-medium education is available, usually as an option, in secondary as well as in primary education. Examples are: Frisian in The Netherlands; Sami in Norway, Sweden, and Finland; Croatian, Hungarian, and Slovene in Austria; Welsh and Gaelic in the United Kingdom.

Elsewhere, particularly in the case of immigrant communities, which are not protected by the Charter, mother-tongue education may be available as a transitional measure in primary school, or it is otherwise left to private initiatives by the community concerned, usually for religious or cultural purposes. In Sweden, the right of all minority children to mother-tongue education was legally established, but proved expensive and difficult to administer to only one individual. A minimum demand from 5 pupils is now required. Most countries strongly encourage long-term residents to acquire the national language, and most adult migrants will wish to do so, except in some cases for elderly relatives and dependents (e.g., Asian women) as well as some expatriate speakers of a supercentral or the hypercentral language. Many countries, including the United Kingdom, require language proficiency as a condition of citizenship. Knowledge of the minority language in other than officially bilingual countries is not widely promoted, except in some cases for social services workers dealing directly with a particular minority group. Provision of court interpreter services and essential documentation in minority languages is the normal situation.

A unique position is occupied by the Romany language, which is the heritage language of an important European community, which is, however, not territorial, because its speakers are nomadic. The language is not well-researched. It exists in a number of spoken forms and has only recently been codified as a written language. Many ethnic Roma do not speak the

language, or do so in a reduced form. Consequently, only Finland, Germany, and Austria, which are not the main centers of Roma populations, appear to make special arrangements for the teaching of Romany. However, a number of Central and Eastern European countries are now concerned to protect and promote Romany language and culture through teacher training materials development and research, and to use Romany as a medium of instruction, as yet on a small scale, in vocational and further education.

Foreign Languages

In foreign language teaching (FLT), the language hierarchy outlined above is particularly clear. In all European countries the teaching of at least one FL is compulsory during secondary education (ages 11–16), except for Ireland, where priority is given to Irish. In the great majority of countries, FLT is compulsory throughout both lower and upper secondary education (for ages 16–18). This has not been the case in the United Kingdom, where modern languages are a minority option in the more specialized sixth form curriculum. The U.K. government has made FL optional from age 14, and about 50% of state comprehensive schools have already dropped FL from the curriculum for Key Stage 4 (ages 14–16). There appears to be a strong correlation between the exercise of the option to continue and social class. In Denmark, FL is not legally required, but 98% of pupils exercise the option.

The teaching of FL in primary education (younger than 11) was discouraged during the 1970s and 1980s following the negative findings of the Burstall Report on the experiment launched in the United Kingdom by the Nuffield Foundation in the 1960s (Burstall *et al.*, 1974). Since 1990, languages in primary schools have rapidly expanded, perhaps most dramatically in Italy. In most countries, the starting age for obligatory FL is between 8 and 11 years of age and is falling. In many countries, FL exists as an option from the start of primary and even in pre-primary education (ages 3–6). U.K. policy aims to compensate for optionality after age 14 by establishing an entitlement for FL teaching in primary education (not obligatory).

The dominant position of English as the hypercentral language is shown, not only by the weaker provision of FL in the anglophone countries, but also by its predominance in FL teaching, learning, and use elsewhere. It is the first and often only FL in all non-anglophone European countries, except Belgium (Dutch and French), Luxembourg (French and German), Romania (French), and Bulgaria and the Baltic Republics (Russian). In most of these cases, special factors operate. In Hungary, Slovakia, and the Czech Republic, German and English are learned as a second language at about the same rate.

A Eurobarometer survey commissioned by the European Commission for the European Year of Languages (2001) showed that strong FL provision enjoyed broad public support. Fully 93% of parents thought it important for their children to learn other European languages, and 71% believed that all Europeans should be competent in one European language other than their mother tongue. Support for two FLs ranged between 20% and 40% except in Greece (55%), Italy (60%), and Luxembourg (80%). About 53% claimed to speak one FL, 47% their mother tongue only. As to bilinguals, 41% of European non-native speakers (70% of students) claimed a command of English. As to the supercentral languages, command of French was claimed by 19% of non-natives, German 10%, Spanish 7%. As the survey did not cover Slav-speaking countries, the position of Russian is hard to assess, particularly because its teaching became identified with the Union of Soviet Socialist Republics' hegemony and collapsed in Eastern Europe after 1990. Of the central languages, only Italian reached 2% among bilinguals.

There are indeed some who resent the dominance of English as 'linguistic imperialism,' an aspect of globalization and U.S. hegemony. This opposition was strong in France, though the role of English in the French education system is no different from that elsewhere. The small but active Esperanto lobby emphasizes its political and cultural neutrality. However, the position of English in the education systems and its role in international life form a self-reinforcing spiral, which is unlikely to change without a radical transformation of the economic, social, and political situation that has produced it.

The Role of European Organizations

The policies of the European agencies strongly favor European cultural and linguistic diversity. Both the Council of Europe and the European Union actively promote measures, particularly in education, to maintain and develop multilingualism. Since the early 1960s, the Council of Europe has provided leadership for the development of FL teaching and learning. Its long-term aims were set out in the preamble to its 1981 Recommendation (81)18 to member governments:

> – that the rich heritage of diverse languages and cultures in Europe is a valuable common resource to be protected and developed, and that a major educational effort is needed to convert that diversity from a barrier to communication into a source of mutual enrichment and understanding;

> – that it is only through a better knowledge of European modern languages that it will be possible to facilitate communication and interaction among Europeans of different mother tongues in order to promote European mobility, mutual understanding and cooperation, and to overcome prejudice and discrimination;
>
> – that member states, when adopting or developing national policies in the field of modern language learning and teaching, may achieve greater convergence at the European level by means of appropriate arrangements for ongoing cooperation and coordination of policies.

In the pursuit of these principles, the Committee of Ministers called upon member governments:

> (F14) To promote the national and international collaboration of government and nongovernment institutions engaged in the development of methods of teaching and evaluation in the field of modern language learning and in the production and use of materials, including institutions engaged in the production and use of multimedia materials.
>
> (F17) To take such steps as are necessary to complete the establishment of an effective European system of information exchange covering all aspects of language learning, teaching, and research, making full use of information technology.

Consequently, the activities of the CDCC, its Committee for Education, and its Modern Languages Section have been concerned to encourage, support, and coordinate the efforts of member governments and nongovernment institutions to improve language learning in accordance with these fundamental principles and in particular the steps that they take to implement the general measures set out in the Appendix to R(82)18:

> A. General Measures
>
> 1. To ensure, as far as possible, that all sections of their populations have access to effective means of acquiring a knowledge of the languages of other member states (or of other communities within their own country) as well as the skills in the use of those languages that will enable them to satisfy their communicative needs and in particular:
>   1.1 to deal with the business of everyday life in another country, and to help foreigners staying in their own country to do so;
>   1.2 to exchange information and ideas with young people and adults who speak a different language and to communicate their thoughts and feelings to them;
>   1.3 to achieve a wider and deeper understanding of the way of life and forms of thought of other peoples and of their cultural heritage.
> 2. To promote, encourage, and support the efforts of teachers and learners at all levels to apply in their own situation the principles of the construction of language-learning systems (as these are progressively developed within the Council of Europe's modern languages programme):
>   2.1 by basing language teaching and learning on the needs, motivations, characteristics, and resources of learners;
>   2.2 by defining worthwhile and realistic objectives as explicitly as possible;
>   2.3 by developing appropriate methods and materials;
>   2.4 by developing suitable forms and instruments for the evaluating of learning programmes.
> 3. To promote research and development programmes leading to the introduction, at all educational levels, of methods and materials best suited to enabling different classes and types of student to acquire a communicative proficiency appropriate to their specific needs.

Its guiding principles, that language learning is for all, is lifelong, and centers on the communicative needs of learners, now enjoy acceptance, largely due to the impact of its Threshold Level concept, first applied to English in 1974 and later updated (van Ek and Trim, 1998) and applied to about 25 European languages. By concentrating on the purposes for which a language is used for communication rather than on its formal structure, a common learning objective was established on which the many agents involved in FL teaching could agree to harmonize their efforts. The success of its practical support for language teaching and teacher training has led to the establishment of the European Modern Languages Centre in Graz, Austria, which conducts an ongoing program of applied research and teacher training workshops, with the particular aim of raising the general level of professional competence to that of acknowledged centers of excellence.

The work of the Language Policy Division of the Council of Europe in Strasbourg has focused on the diversification of language learning and on the innovative concept of *plurilingualism*, related to but distinct from multilingualism. Plurilingualism asserts that, in the course of their experience, language users gradually build up a global communicative competence to which all language learning and use has a contribution to make. Languages are not held separate in the mind. They interact and together make new learning easier and can all be drawn upon for communication as the situation requires. Learning and teaching do not have to concentrate on perfect performance in a single language, or even two, but on enlarging the overall language competence and the effectiveness of its use. This concept of plurilingualism clearly implies a radical rethinking of language pedagogy, which is still at an early stage. To

assist, the Council of Europe has developed two new instruments. The *European language portfolio* enables learners to document the whole of their language learning experience, which may or may not lead to formal qualifications. The *Common European framework of reference for languages: learning, teaching, assessment* (CEFR) provides categories for all aspects of language use and the different kinds of competence that a language user has to develop and call upon for communication. It also provides descriptors for six levels of proficiency, A1 to C2. CEFR is already widely used across the Continent, partly for teacher training (Little and Perclová, 2001), but mainly to make the level of national qualifications more transparent and comparable. The preparation of a manual of guidance was started in 2004 by an international working party. A series of national language policy profiles is also underway. Reports have been produced by authorities in Hungary, Lithuania, Norway, and Slovenia.

The engagement of the European Union in language teaching policy is more recent and different in character. Whereas the Council of Europe is an intergovernmental organization, working toward common action through consensus and recommendations, the European Union is a supranational body, able to issue directives to member states, who are the more watchful for the limits of its powers. Until the Maastricht Treaty, the authority of the EU was limited to vocational (including teacher) training. Since Maastricht, it has been able to operate in the general educational field, but only in accordance with the principle of subsidiarity. It cannot determine policy, but acts in support of member states' government action where action is best organized on a European level. The broad lines of such action were outlined in the *White paper on teaching and learning: towards the learning society* (European Commission, 1995). Action IV: "proficiency in three Community languages" gave priority to supporting the learning by all Europeans of two EU languages in addition to the mother tongue. Because the first FL is almost always English, the requirement for two FL seems designed to reduce the dominance of English. However, the Commission was obliged to accept, though it deplored, the reluctance of member governments to give wholehearted support to a compulsory second FL by inserting the words *wherever possible* into the substantive resolution of the Council of Ministers. The Eurobarometer quoted above shows that public opinion shares the reservations of governments, and it seems likely that in most EU member states, one FL, usually English, will remain the universal obligation, with second and third FLs reserved for more able students, specialists, or in clearly multilingual situations, until such time as the plurilingual approach is more widely understood and accepted.

The main instruments used by the EU to support and influence FL teaching in Europe are its funding programs. The SOCRATES Programme for education in Europe comprises a number of actions relevant to FLs. COMENIUS focuses on the first phase of education, from preschool and primary to secondary school. It is addressed to a broad range of interests: pupils, teachers and other education staff, as well as local authorities, parents' associations, and relevant nongovernment organizations—in short, all those whose decisions determine policies and their implementation. By supporting school networks and partnerships and staff training, it aims to enhance the quality of teaching, strengthen the European dimension, and promote language learning and mobility. The issues emphasized include: learning in a multicultural framework, seen as the cornerstone of European citizenship, support for disadvantaged groups, and countering underachievement. ERASMUS concerns higher education, supporting student and teacher exchanges, joint development of study programs, thematic networks among departments and faculties across Europe, and the development of a European credit transfer system. GRUNDTVIG completes the coverage of lifelong education by supporting activities designed to promote European cooperation in improving the availability, accessibility, and quality of adult education. LINGUA is directly concerned with the promotion of language teaching and learning. The Lingua Action is designed to encourage and support linguistic diversity throughout the EU, to improve the quality of language teaching and learning, and to promote access to lifelong learning opportunities appropriate to each individual's needs. The action is divided into two parts: LINGUA 1 is intended to raise citizens' awareness of the Union's multilingual wealth, to encourage people to learn languages throughout their lifetime, and to improve access to FL learning resources across Europe as well as to develop and disseminate innovative techniques and good practice in language teaching. LINGUA 2 aims to ensure that a sufficiently wide range of language learning tools is available to language learners.

Support for language learning is also forthcoming for public and/or private bodies under the EC's vocational training program LEONARDO da VINCI. In addition to international placements, exchanges, and networks for the exchange of experience and good practice, Language Competences Projects' aim is to enhance multilingual and multicultural communication in training and working environments, e.g., by developing training tools and methods for training

language examiners. The TEMPUS program, which was initiated in 1990 to respond to the needs for higher education (HE) reform in Central and Eastern European countries following the fall of the Berlin Wall, encourages structural cooperation by forming consortia joining higher education institutions in EU member states with their counterparts in a wider range of non-EU states. With varying arrangements for financing, co-financing or self-financing, the consortia implement Joint European Projects with clearly defined objectives for up to three years. TEMPUS also provides Individual Mobility Grants.

The financial resources made available under these programs are large. They are eagerly sought and have been instrumental in bringing increasing numbers of schools and university departments to see their activities in a collaborative international context. Lacking such resources, the work of the Council of Europe has had to rely on its intellectual content persuading institutions and individuals in member states that their own interests are best served by investing some part of their own resources in international projects, which is perhaps a useful discipline.

A further consequence of financial disparity has been the internal language policies of the two organizations. Since its foundation, the Council has had just two official languages, English and French. In principle, the EU accords equal status to the national languages of all its members. Strictly applied, this would mean a huge operation of interpretation and translation for every event. In practice, that apparatus is deployed for sessions of the European Parliament and a number of formal occasions. Informal meetings are usually conducted, by agreement, in English and French. The siting of major institutions in French-speaking cities means that French is generally the predominant house language, while many outside delegates and experts use English.

Conclusion

Europe is, and will remain, a region of high linguistic diversity. Language policy at all its many levels is, and will continue to be, torn between the pressures of globalization and ever more efficient means of communication on the one hand, and strong feelings of national and community identity on the other. Political decision-makers will have to assess the relative strength of each pressure in specific situations, to find solutions that avoid or at least contain conflict. Much will depend on the success of education systems in producing a well-informed citizenry, free from mistrust and xenophobia, able and willing to form good working and warm personal relations with fellow human beings with different linguistic and cultural backgrounds.

See also: Applied Linguistics in Europe; Education in a Multilingual Society; Foreign Language Teaching Policy; Immigrant Language Education; Language Policies: Policies on Language in Europe; Minority Languages: Education.

Bibliography

Ager D (2001). *Motivation in language planning and language policy.* Clevedon: Multilingual Matters.

Beacco J-C & Byram M (2002). *Guide for the development of language education policies in Europe: from linguistic diversity to plurilingual education. Main version, Draft 1.* Strasbourg: Council of Europe Language Policy Division.

Burstall C, Jamieson M, Cohen S & Hargreaves M (1974). *Primary French in the balance.* Windsor: NFER Publishing Company.

Council of Europe (1992). *European charter for regional or minority languages. European Treaty series 148.* Strasbourg: Council of Europe.

Council of Europe (2002). *Common European framework of reference for languages: learning, teaching and assessment.* Cambridge: Cambridge University Press.

Council of Europe Committee of Ministers (1982). *Recommendation no. R (82)18 of the committee of ministers to member states concerning modern languages.* Strasbourg: Council of Europe.

De Swaan A (2001). *Words of the world.* Cambridge: Polity Press.

Dickson P & Cumming A (1996). *Profile of language education in 25 countries.* Slough: National Foundation for Educational Research.

European Commission (1995). *White paper on teaching and learning: towards the learning society.* Luxembourg: European Commission.

European Commission (2000). *Guide to programmes and actions.* Luxembourg: European Commission.

EURYDICE (2001). *Foreign language teaching in schools in Europe.* Brussels: Eurydice European Unit.

EURYDICE & Eurostat (2002). *Key data on education in the European Union.* Luxembourg: European Commission.

Extra G & Gorter D (2001). *The other languages of Europe: demographic, sociolinguistic and educational perspectives.* Clevedon: Multilingual Matters.

Kaplan R B & Baldauf R B, Jr. (1997). *Language planning: from practice to theory.* Clevedon: Multilingual Matters.

Little D & Perclová R (2001). *European language portfolio: guide for teachers and teacher trainers.* Strasbourg: Council of Europe.

Trim J L M (1994). 'Some factors influencing national foreign language policymaking in Europe.' In Lambert R D (ed.) *Language planning around the world: contexts and systemic change.* Washington, D.C.: National Foreign Language Center.

Trim J L M (2002). 'Foreign language policies in Europe, with special reference to the role of the council of Europe and the European Union.' In Baker S J (ed.) (2001). *Language policy: lessons from global models.* Monterey, CA: Monterey Institute of International Studies.
Universal declaration of linguistic rights (1998). Barcelona: Follow-up Committee to the Universal Declaration of Linguistic Rights.
Van Ek J & Trim J L M (1998). *Threshold 1990.* Cambridge: Cambridge University Press.
Van Haecht A (1996). *The division of responsibilities at national, regional and local levels in the education systems of twenty-three European countries.* Strasbourg: Council of Europe.

Language Education Policy in Latin America

K King, Georgetown University, Washington, DC, USA

Whereas language planning concerns the decisions people make about languages and their uses in society, language education policy is the means by which those decisions are implemented "in the specific contexts of schools and universities in relation to home languages" (Shohamy, 2003: 279). In Latin America, language education policy has long been the primary channel for implementing national language policy, as education through the Spanish language has been the principal means of establishing a policy of linguistic and cultural assimilation for the indigenous population. In more recent years, progressive educational reforms that promote intercultural bilingual education have been implemented in conjunction with national language policies that recognize and, in some cases, officialize indigenous languages. Language education policy thus serves as the point of contact between abstract national policy on the one hand and the lives of regular citizens on the other.

Most of Latin America is characterized by long-standing economic and cultural gaps between three societal sectors: the indigenous population (also known as Indians or originary peoples); individuals who claim mixed Spanish and indigenous heritage (often known as *mestizos*); and *blancos* (whites), who claim or are perceived as holding Spanish or European heritage. Although estimates vary widely across Latin America (see Albó, 2004, for some reasons why), there are approximately 40 million indigenous people in Latin America, who speak one or more of roughly 400 indigenous languages, constituting approximately 10% of the total population (HRDOP, 1993). Demographics and the nature of the relationships between these populations of course vary across Latin American nations. For instance, in Brazil, as many as 150 to 170 indigenous languages are spoken by approximately one-quarter of a million Indians (but who represent less than 1% of the total population of 182 million) (World Factbook, 2003a). In Argentina, estimates concerning the indigenous population range from 450,000 to 1.5 million, representing only 1 to 4% of the total population of approximately 36 million (HRDC, 2001). In Mexico, in turn, there are 62 indigenous languages spoken by approximately 10 million people (representing approximately 10% of the population) (INEGI, 2000). In Bolivia, 70% (of the total population of 8 million) are estimated to be indigenous (Albó, 1995; ETARE, 1993) and over 30 ethnolinguistic groups have been identified, the largest of which are North and South Bolivian Quechua, Central Aymara (1.6 million), Eastern and Western Guaraní (60 000) (MDH, 1995).

Yet despite these differences and with some notable exceptions (for example, during the precolonial and early colonial period in Central Mexico, a variety of Náhautl was the principal language of instruction and literacy, Francis, 2002), the dominant narrative in most of Latin America has been one of Spanish domination and long-standing inequality, with indigenous sectors suffering from limited legal rights, restricted access to education, healthcare, and other social services, and less than full participation in their countries' political and economic systems. Language education policy both reflects and reinforces these inequalities, as for many decades in many regions of Latin America, indigenous students were, at best, denied education in a language that they could comprehend and, at worst, harshly punished for speaking their home languages in the classroom.

Very broadly speaking, Latin American language education policy falls into two broad types. The first set of policies is aimed at students who are monolingual speakers of Spanish (or in Brazil, Portuguese) acquiring English (or less commonly, French or another European language) as a foreign language. In many regions, these programs tend to be designed for – and in many instances are only available to – students from middle- to upper-class backgrounds and are created as a means to enrich their educational

Trim J L M (2002). 'Foreign language policies in Europe, with special reference to the role of the council of Europe and the European Union.' In Baker S J (ed.) (2001). *Language policy: lessons from global models.* Monterey, CA: Monterey Institute of International Studies.
Universal declaration of linguistic rights (1998). Barcelona: Follow-up Committee to the Universal Declaration of Linguistic Rights.
Van Ek J & Trim J L M (1998). *Threshold 1990.* Cambridge: Cambridge University Press.
Van Haecht A (1996). *The division of responsibilities at national, regional and local levels in the education systems of twenty-three European countries.* Strasbourg: Council of Europe.

Language Education Policy in Latin America

K King, Georgetown University, Washington, DC, USA

Whereas language planning concerns the decisions people make about languages and their uses in society, language education policy is the means by which those decisions are implemented "in the specific contexts of schools and universities in relation to home languages" (Shohamy, 2003: 279). In Latin America, language education policy has long been the primary channel for implementing national language policy, as education through the Spanish language has been the principal means of establishing a policy of linguistic and cultural assimilation for the indigenous population. In more recent years, progressive educational reforms that promote intercultural bilingual education have been implemented in conjunction with national language policies that recognize and, in some cases, officialize indigenous languages. Language education policy thus serves as the point of contact between abstract national policy on the one hand and the lives of regular citizens on the other.

Most of Latin America is characterized by long-standing economic and cultural gaps between three societal sectors: the indigenous population (also known as Indians or originary peoples); individuals who claim mixed Spanish and indigenous heritage (often known as *mestizos*); and *blancos* (whites), who claim or are perceived as holding Spanish or European heritage. Although estimates vary widely across Latin America (see Albó, 2004, for some reasons why), there are approximately 40 million indigenous people in Latin America, who speak one or more of roughly 400 indigenous languages, constituting approximately 10% of the total population (HRDOP, 1993). Demographics and the nature of the relationships between these populations of course vary across Latin American nations. For instance, in Brazil, as many as 150 to 170 indigenous languages are spoken by approximately one-quarter of a million Indians (but who represent less than 1% of the total population of 182 million) (World Factbook, 2003a). In Argentina, estimates concerning the indigenous population range from 450,000 to 1.5 million, representing only 1 to 4% of the total population of approximately 36 million (HRDC, 2001). In Mexico, in turn, there are 62 indigenous languages spoken by approximately 10 million people (representing approximately 10% of the population) (INEGI, 2000). In Bolivia, 70% (of the total population of 8 million) are estimated to be indigenous (Albó, 1995; ETARE, 1993) and over 30 ethnolinguistic groups have been identified, the largest of which are North and South Bolivian Quechua, Central Aymara (1.6 million), Eastern and Western Guaraní (60 000) (MDH, 1995).

Yet despite these differences and with some notable exceptions (for example, during the precolonial and early colonial period in Central Mexico, a variety of Náhautl was the principal language of instruction and literacy, Francis, 2002), the dominant narrative in most of Latin America has been one of Spanish domination and long-standing inequality, with indigenous sectors suffering from limited legal rights, restricted access to education, healthcare, and other social services, and less than full participation in their countries' political and economic systems. Language education policy both reflects and reinforces these inequalities, as for many decades in many regions of Latin America, indigenous students were, at best, denied education in a language that they could comprehend and, at worst, harshly punished for speaking their home languages in the classroom.

Very broadly speaking, Latin American language education policy falls into two broad types. The first set of policies is aimed at students who are monolingual speakers of Spanish (or in Brazil, Portuguese) acquiring English (or less commonly, French or another European language) as a foreign language. In many regions, these programs tend to be designed for – and in many instances are only available to – students from middle- to upper-class backgrounds and are created as a means to enrich their educational

and social opportunities. Enrichment programs such as these, which promote an additive type of bilingualism, have sometimes been labeled 'elitist' bilingual education (cf. Hornberger, 1991; de Mejia, 2002). See Friedrich and Berns (2003) for a review of literature on English in South America.

The second set of language education policies, which will be the focus of the remainder of this article, is aimed at students who speak an indigenous language and who need to acquire Spanish (or Brazilian Portuguese) as a second language. Such programs are designed as vehicles for providing meaningful instruction and literacy training in the students' first language, prior to (or concomitant with) their transition to Spanish (or Portuguese). These students tend to come from economically and socially marginalized groups within the national context. Advocates of such programs argue that the use of the students' first language provides them with a greater chance of engaging with the school curriculum, developing literacy skills, and eventually, participating on more equal terms in the wider national society. Yet these programs tend to be oriented overwhelmingly toward transition to the official language of the country and hence to promote a subtractive form of bilingualism.

Both types of language education policies have received greater attention in recent decades: the former as the result of globalization and the (perceived and real) importance of English on the global stage (Gunaratne, 2003) and the latter in conjunction with greater recognition of individual human rights and the progressive political reforms in many Latin American nations (Luykx, 2000, 2004; Skutnabb-Kangas, 2000). As a result, intercultural bilingual education has become much more widely recognized; indeed, "almost every Latin American country is developing this type of program, with the government, organized social groups, or both functioning as its source" (Godenzzi, 2003: 295).

Yet despite significant policy shifts, the long-term social inequalities that many language education policies were designed to address have yet to be substantially altered in the region, in part due to the significant lag in implementation of these measures. What von Gleich observed a decade ago for Peru remains true in much of the region; that is, "the gap between even the modest legal status [for the indigenous language] and implementation is deep" (1992: 62). The major challenges to full implementation of language education policy that meaningfully uses indigenous languages in classrooms include the dearth of qualified bilingual teachers, questions of language standardization, divergent interpretations of 'bilingual' and 'intercultural,' and complications surrounding decentralization and local authority over indigenous schools (see King and Benson, 2003, for a full discussion). Each of these is briefly discussed in turn below.

A central issue in expanding the use of indigenous languages in education is the cultivation of qualified bilingual teachers (see, for instance, Abram, 1989). The lack of such teachers is not surprising considering the traditionally low status and levels of formal education associated with indigenous sectors, as well as the limited salaries earned by rural teachers in general. Even where basic training and recruitment issues are addressed, ideological forces potentially undermine the use of indigenous languages in the classroom. As Dauenhauer and Dauenhauer (1998) describe, there is often disparity between expressed ideals and actual support for indigenous languages, which, in their view, often results from deeply embedded ideologies concerning the language. For instance, indigenous teachers might sharply recall their own punishment and embarrassment for using their language in school.

A second important challenge to implementing some language education policies stems from complications with the standardization of indigenous languages and development of pedagogical materials. The example of Quichua in Ecuador is illustrative. The process of standardizing the different varieties of Quichua was formally initiated in 1981 when representatives of the different Ecuadorian varieties of Quichua agreed on a unified variety, known as Quichua Unificado. These decisions were codified in subsequent dictionaries and grammars (CONAIE, 1990; MEC, 1982), constituted a major step toward the standardization of language, and were intended to facilitate the development of Quichua materials, as well as contribute to the maintenance and even revitalization of the language. In practice, however, these aims have proved elusive. For instance, King (1999, 2000) reports that, as Quichua materials have been introduced into Quichua–Spanish bilingual communities in the southern highlands, two distinct Quichua varieties have emerged. The Quichua pedagogical materials promote the nationally standardized variety (Quichua Unificado), in contrast to what is commonly referred to as Quichua auténtico, spoken by elderly and rural dwellers. Because children and young adults studying Quichua Unificado have not mastered the phonological system or the lexicon of the local variety, they learn not only to read but to speak Quichua Unificado. Although the varieties are mutually intelligible to most, there are clashes and gaps in communication between the older and younger Quichua speakers and such tensions undermine Quichua use and exacerbate generational and social divisions.

Cases such as these have led some to question the assumption that standardization of the written form is essential to expanded use of the language. Luykx (2000, 2004), for instance, maintains that in practice the emphasis on standardization tends to shift resources away from those domains that are the indigenous language's stronghold and toward those where its disadvantage relative to Spanish is greatest, to prioritize concerns of linguists over those of most native speakers, and to link dialectal variation with status and class hierarchies. The result is that standardization tends to undermine meaningful use of indigenous languages in communities and schools. Although standardization has long been viewed as an important means to improving status and prestige (Hornberger and King, 1996; Stewart, 1968), it seems appropriate to question whether it is essential to the establishment of successful bilingual education programs.

A third issue concerns how 'bilingual' and 'intercultural' are operationalized. Not unlike early formulations of bilingual education policy in the United States (Crawford, 1999), legislation in many countries provides few specifics concerning how bilingual education is defined and how Spanish and the indigenous language(s) are balanced within and across grades. Likewise, intercultural education is also open to interpretation, despite the fact that "bilingual intercultural education" is the standard term in use in much of the region. Krainer (1996) defines intercultural education as that which affirms one's social and conceptual universe, but also permits the selective and critical appropriation of cultural elements of other groups. Haboud (unpublished) argues that interculturalism refers to relations and interactions between two separate cultures, as well as the political task of constructing an egalitarian society. However, as frequently noted, interculturalism has often been treated as an issue that is exclusive to the indigenous population (Haboud, 2000; Hornberger, 2000; Hornberger and Coronel-Molina, 2004).

A final issue relates to the delegation of authority to indigenous groups to control their own schools; in many regions, this is part of the move to decentralize and democratize government administration. In the case of the Bolivian Reform, despite the democratic aims of decentralization and popular participation, the Ministry of Education maintains a tight hold on most aspects of the transformation of traditional schools to Reform schools. One avenue for indigenous participation is through the four CEPOs (*Consejos Educativos de los Pueblos Originarios*), which were formed during the 1990s to represent Quechua, Aymara, Guaraní, and speakers of various Amazonian languages in national educational concerns. The CEPOs have been instrumental in garnering community support for bilingual intercultural education, promoting linguistic development and training of teachers and writers, and organizing the Reform-mandated school boards at the school, district, and departmental levels. However, their relationship to the Ministry of Education is undefined; their autonomous status simultaneously protects them from government politics and leaves them at the mercy of international donors for salaries and support, and the Ministry is not required to incorporate their advice (Albó, 2000; Moya, 1999).

It is also important to note that greater local control of indigenous schooling and increased numbers of indigenous teachers do not necessarily or immediately translate into substantial use of indigenous languages in the classroom. King's (2000) study of two communities in southern Ecuador suggests that although Quichua is employed regularly in schools for symbolic purposes (e.g., greetings, taking attendance), it is taught only occasionally and can in no way be considered a medium of instruction. This frequent symbolic use of Quichua in school, in contrast to its total absence prior to the reforms, obfuscates the limited academic and communicative use of the language.

As evident here, some of the shifts in language policy have been largely rhetorical as they remain without programs or funding to implement any real change. It is also true, however, that "gestures that, at first, seemed merely symbolic open new spaces of negotiation between indigenous organizations and different state sectors" (Godenzzi, 2003: 295). Indigenous groups and individuals are increasingly taking on leadership roles and administrative responsibilities in reform efforts. Many indigenous languages have written forms and are now used in literacy and basic schooling. Increasing numbers of indigenous people are becoming literate in their mother tongues or learning the mother tongues of their grandparents, and the presence of indigenous languages in a once exclusively Spanish domain confers prestige on both the languages and their users. Although the gaps between language education policy and practice remain frustratingly wide, many important small steps have been made and continued progress into this new, uncharted territory seems likely to follow.

See also: Bilingual Education; Language Planning and Policy: Models; Teaching of Minority Languages.

Bibliography

Abram M (1989). 'El papel de las organizaciones indigenas en la educación.' In Lopez E & Moya R (eds.) *Pueblos indios, estados, y educación*. Lima: PEB/PEBI/PERA. 309–326.

Albó X (1995). *Bolivia plurilingüe. Guía para planificadores y educadores*. La Paz, Bolivia: UNICEF y CIPCA.

Albó X (2000). *Evaluación del componente 'Apoyo a la educación intercultural bilingüe en tierras bajas de Bolivia.'* La Paz: CIPCA.

Albó X (2004). 'El Quechua visto desde una perspectiva Boliviana.' *International Journal of the Sociology of Language 167,* 119–130.

CONAIE (Confederación de Nacionalidades Indígenas del Ecuador) (1990). *Ñucanchic shimi 1*. Quito: MEC/DINEIIB/CONAIE.

Crawford J (1999). *Bilingual education: History, politics, theory, and practice*. Los Angeles: Bilingual Education Services.

Dauenhauer N M & Dauenhauer R (1998). 'Technical, emotional, and ideological issues in reversing language shift: Examples from Southeast Alaska.' In Grenoble L A & Whaley L J (eds.) *Endangered languages: Current issues and future prospects*. Cambridge, UK: Cambridge University Press. 57–116.

de Mejia A M (2002). *Power, prestige, and bilingualism: International perspectives on elite bilingual education*. Clevedon, UK: Multilingual Matters.

ETARE (Equipo Técnico de Apoyo a la Reforma Educativa) (1993). *Reforma Educativa. Propuesta. Cuadernos de la Reforma Educativa*. La Paz: MDH/SNE.

Francis N (2002). 'Models of bilingual proficiency for a pluralistic language policy in Latin America.' *Communication and Cognition 35(1/2),* 101–122.

Friedrich P & Berns M (2003). 'English in South America: Selected bibliographic resources.' *World Englishes 22(2),* 211–214.

Godenzzi J G (2003). 'Perspectives from Latin American intercultural and bilingual education.' *Modern Language Journal 87,* 294–296.

Gunaratne S A (2003). 'Proto-Indo-European expansion, rise of English, and the international language order: A humanocenteric analysis.' *International Journal of the Sociology of Language 164,* 1–32.

Haxton M L, Miller T & Cousins K (1995). 'Indigenous peoples in Paraguay.' Retrieved March 29, 2004 from http://www.cidcm.umd.edu/inscr/mar/data/indpar.htm.

Hornberger N H (1991). 'Extending enrichment bilingual education: Revisiting typologies and redirecting policy.' In Garcia O (ed.) *Bilingual education: Foccushrift in honor of Joshua A. Fishman,* 1. Philadelphia: John Benjamins. 215–234.

Hornberger N H (2000). 'Bilingual education policy and practice in the Andes: Ideological paradox and intercultural possibility.' *Anthropology and Education Quarterly 31(2),* 173–201.

Hornberger N H & Coronel-Molina S (2004). 'Quechua language shift, maintenance and revitalization in the Andes: The case for language planning.' *International Journal of the Sociology of Language 167,* 9–67.

Hornberger N H & King K A (1996). 'Language revitalization in the Andes: Can the schools reverse language shift?' *Journal of Multilingual and Multicultural Development 17(6),* 427–441.

HRDC (Human Rights Documentation Center) (2001). *'Executive summary, racial discrimination: The record of Argentina.'* Retrieved March 29, 2004 from. http://academic.udayton.edu/race/06hrights/GeoRegions/SouthAmerica/argentina01.htm.

HRDOP (Human Resources Development and Operations Policy) (1993). *'Indigenous people in Latin America.'* Retrieved March 29, 2004 from http://www.worldbank.org/html/extdr/hnp/hddflash/hcnote/hrn007.html.

INEGI (Instituto Nacional de Estadística, Geografía e Informática) (2000). *'La diversidad cultural en México, Los pueblos indígenas y sus 62 idiomas.'* Retrieved March 29, 2004 from http://www.elbalero.gob.mx/kids/about/html/indigenous/lenguas.html.

King K A (1999). 'Inspecting the unexpected: Language corpus and status shifts as aspects of language revitalization.' *Language Planning and Language Problems 23(2),* 109–132.

King K A (2000). *Language revitalisation processes and prospects: Quichua in the Ecuadorian Andes*. Clevedon, UK: Multilingual Matters.

King K A & Benson C (2003). 'Indigenous language education in Bolivia and Ecuador: Contexts, changes, and challenges.' In Tollefson J W & Tsui A B M (eds.) *Medium of instruction policies: Whose agenda? Which agenda?*. Mahwah, NJ: Lawrence Erlbaum Associates. 241–261.

Krainer A (1996). *Educacion intercultural bilingüe en el Ecuador.* Quito: Ediciones Abya-Yala.

Krauss M (1998). 'The condition of Native North American languages: The need for realistic assessment and action.' *International Journal of the Sociology of Language 132,* 9–21.

Luykx A (2000). 'Diversity in the new world order: State language policies and the internationalization of Quechua.' Paper presented at the 2nd Spencer Early Career Institute in Anthropology and Education: Globalization and Education. May 10–13, 2000, Chicago, IL.

Luykx A (2004). 'The future of Quechua and the Quechua of the future: Language ideologies and language planning in Bolivia.' *International Journal of the Sociology of Language 167,* 147–158.

MDH (Ministerio de Desarrollo Humano) (1995). *Secretaria de Educación, UNSTP (1995). Nuevos programas de estudio*. La Paz: MDH.

MEC (Ministerio de Educacion y Cultura) (1982). *Caimi Ñucanchic Shimiyuc-Panca*. Quito: Ministerio de Educación y Cultura, Pontificia Universidad Católica del Ecuador.

Moya R (1999). *Informe para el gobierno de Suecia sobre la evaluatción de la Reform Educativa de la gestión de 1999. Componente de Educación Intercultural Bilingüe*. La Paz: Sida.

Shohamy E (2003). 'Implications of language education policies for language study in schools and universities.' *Modern Language Journal 87,* 278–286.

Skutnabb-Kangas T (2000). *Linguistic genocide in education—or worldwide diversity and human rights?* Mahwah, NJ: Lawrence Erlbaum Associates.

Stewart W (1968). 'A sociolinguistic typology for describing national multilingualism.' In Fishman J A (ed.) *Readings in the sociology of language.* The Hague: Mouton. 531–545.

von Gleich U (1992). 'Changes in the status and function of Quechua.' In Ammon U & Hellinger M (eds.) *Status changes of languages.* New York: Walter de Gruyter. 43–64.

World Factbook (2003a). 'Brazil.' Retrieved March 29, 2004 from http://www.cia.gov/cia/publications/factbook/geos/br.html.

World Factbook (2003b). 'Paraguay.' Retrieved March 29, 2004 from http://www.cia.gov/cia/publications/factbook/geos/pa.html.

Language Education Policy in Soviet Successor States

L Grenoble, Dartmouth College, Hanover, NH, USA

The Soviet Legacy

From the time it was founded in 1917, the Soviet Union was a multilingual and multiethnic state. The 1989 Soviet census cites approximately 130 ethnic groups (including indigenous and immigrant populations) speaking a total of 150 languages (although linguists place the count closer to 200). Russians comprise the largest single ethnic group, constituting of a little more than 50% of the total population (or slightly more than 145 000 000) people in the 1989 census. The Russian language was used by the largest number of speakers: 81% of the population identified itself as fluent in Russian, either as a first or second language. By the time of the breakup of the Soviet Union in 1991, Russian was firmly established as the lingua franca for the entire country.

Fifteen states were created with the dissolution of the Soviet Union: the Russian Federation (Russia); the Baltic States: Estonia, Latvia, and Lithuania; three in the Caucasus: Armenia, Georgia, and Azerbaijan; five Republics in Central Asia: Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan; and Belarus, Moldova, and Ukraine. There is great variation among these successor states in terms of numbers of languages spoken, speaker density, levels of multilingualism, and official language policy.

Historical Background

Although its early language policies promoted the right of national languages in government and education, by the post World War II period, the Soviet-government had firmly established a language policy that favored the use of Russian over all other languages in its territory. By the late 1980s, the expansion of Russian at the expense of the titular languages caused a backlash from all regions in the USSR, resulting in the introduction of language laws in most of the Soviet Republics. The movement began in the Baltics, with the Estonian SSR on January 18, 1989; the Lithuanian SSR on January 25; and the Latvian SSR on May 5. It then spread: the Tajik SSR on July 22; the Moldovan SSR on September 1; the Kazakh SSR on September 22; the Kirghiz SSR on September 23; the Uzbek SSR on October 21, and the Ukrainian SSR on October 28; finally, in 1990, the Belorussian SSR on January 26 and the Turkmen SSR on May 24. Only four Republics failed to adopt new language laws in this time period: the Russian SFSR and all three Republics in the Caucasus (the Armenian SSR, Azerbaijan SSR, and the Georgian SSR). The latter three did not adopt such laws because they would have been redundant: the basic principles governing language use had been specified since the ratification of their new Constitutions in April 1978. The Russian SFSR did not submit a language law until October 1991, when the Soviet breakup was well under way.

Although the laws differ in many details, the key point of each was the identification of the titular language as the official state language. This was intended to foster its development and use, and to guarantee its primacy in the government and legal system, education, and the media. Most laws also give their citizens the right to use Russian, or another regional language, in these contexts. Furthermore, most required knowledge of both the state language and Russian for employees in state offices and organizations.

Almost all the language laws specified a timeframe for the transition from Russian to the newly declared State language, and all but the Kirghiz and Tajik laws provided specific deadlines. The deadlines varied from three to eight years from the initiation of each individual law. They were optimistic but unrealistic; as early as summer 1991 in the Kirghiz SSR (*see* **Kirghiz**), the official deadline was changed from April 1993 until the year 2000. Delays were

Shohamy E (2003). 'Implications of language education policies for language study in schools and universities.' *Modern Language Journal 87,* 278–286.

Skutnabb-Kangas T (2000). *Linguistic genocide in education—or worldwide diversity and human rights?* Mahwah, NJ: Lawrence Erlbaum Associates.

Stewart W (1968). 'A sociolinguistic typology for describing national multilingualism.' In Fishman J A (ed.) *Readings in the sociology of language.* The Hague: Mouton. 531–545.

von Gleich U (1992). 'Changes in the status and function of Quechua.' In Ammon U & Hellinger M (eds.) *Status changes of languages.* New York: Walter de Gruyter. 43–64.

World Factbook (2003a). 'Brazil.' Retrieved March 29, 2004 from http://www.cia.gov/cia/publications/factbook/geos/br.html.

World Factbook (2003b). 'Paraguay.' Retrieved March 29, 2004 from http://www.cia.gov/cia/publications/factbook/geos/pa.html.

Language Education Policy in Soviet Successor States

L Grenoble, Dartmouth College, Hanover, NH, USA

The Soviet Legacy

From the time it was founded in 1917, the Soviet Union was a multilingual and multiethnic state. The 1989 Soviet census cites approximately 130 ethnic groups (including indigenous and immigrant populations) speaking a total of 150 languages (although linguists place the count closer to 200). Russians comprise the largest single ethnic group, constituting of a little more than 50% of the total population (or slightly more than 145 000 000) people in the 1989 census. The Russian language was used by the largest number of speakers: 81% of the population identified itself as fluent in Russian, either as a first or second language. By the time of the breakup of the Soviet Union in 1991, Russian was firmly established as the lingua franca for the entire country.

Fifteen states were created with the dissolution of the Soviet Union: the Russian Federation (Russia); the Baltic States: Estonia, Latvia, and Lithuania; three in the Caucasus: Armenia, Georgia, and Azerbaijan; five Republics in Central Asia: Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan; and Belarus, Moldova, and Ukraine. There is great variation among these successor states in terms of numbers of languages spoken, speaker density, levels of multilingualism, and official language policy.

Historical Background

Although its early language policies promoted the right of national languages in government and education, by the post World War II period, the Soviet-government had firmly established a language policy that favored the use of Russian over all other languages in its territory. By the late 1980s, the expansion of Russian at the expense of the titular languages caused a backlash from all regions in the USSR, resulting in the introduction of language laws in most of the Soviet Republics. The movement began in the Baltics, with the Estonian SSR on January 18, 1989; the Lithuanian SSR on January 25; and the Latvian SSR on May 5. It then spread: the Tajik SSR on July 22; the Moldovan SSR on September 1; the Kazakh SSR on September 22; the Kirghiz SSR on September 23; the Uzbek SSR on October 21, and the Ukrainian SSR on October 28; finally, in 1990, the Belorussian SSR on January 26 and the Turkmen SSR on May 24. Only four Republics failed to adopt new language laws in this time period: the Russian SFSR and all three Republics in the Caucasus (the Armenian SSR, Azerbaijan SSR, and the Georgian SSR). The latter three did not adopt such laws because they would have been redundant: the basic principles governing language use had been specified since the ratification of their new Constitutions in April 1978. The Russian SFSR did not submit a language law until October 1991, when the Soviet breakup was well under way.

Although the laws differ in many details, the key point of each was the identification of the titular language as the official state language. This was intended to foster its development and use, and to guarantee its primacy in the government and legal system, education, and the media. Most laws also give their citizens the right to use Russian, or another regional language, in these contexts. Furthermore, most required knowledge of both the state language and Russian for employees in state offices and organizations.

Almost all the language laws specified a timeframe for the transition from Russian to the newly declared State language, and all but the Kirghiz and Tajik laws provided specific deadlines. The deadlines varied from three to eight years from the initiation of each individual law. They were optimistic but unrealistic; as early as summer 1991 in the Kirghiz SSR (*see* **Kirghiz**), the official deadline was changed from April 1993 until the year 2000. Delays were

inevitable because many national languages had not been developed for administrative, technical, and educational purposes; only Russian had been used in these domains for many decades. Moreover, ethnic Russians living in each Republic resisted these changes, and in some areas local people had ceased using the titular languages in favor of Russian.

The Russian Federation

On October 25, 1991, the Supreme Soviet of the USSR ratified the "Law of the languages of the peoples of the Russian Soviet Federated Socialist Republic." This law confirmed Russian as the official state language, leaving the status of other languages to the discretion of the individual republics. In 1993, the Russian Federation (*see* **Russian Federation: Language Situation**) reratified this law with only minor changes.

With the collapse of the Soviet Union, political reconfiguration of the territories remaining in the Russian Federation resulted in the creation of a new set of republics, many being formed from previous Autonomous Soviet Socialist Republics (e.g., the Chuvash Republic or the Republic of Tatarstan). Of the 21 Republics in the Russian Federation at the end of the 20th century, 15 had instituted language laws to grant their titular language equal official status with Russian. Three of these laws (in the Chuvash, Tuvan, and Kalmyk Republics) were passed before the downfall of the Soviet state, with the Chuvash law in October 1990, the Tuvan law in December 1990, and the Kalmyk law in January 1991. In addition, the Constitutions of these Republics include language-specific articles; in most cases they grant official status to the titular language and Russian.

At present, a number of languages have some sort of official status in the Russian Federation. Russian is the state language for the entire country. The titular languages of the Republics within the Russian Federation are: Adyge, Altai, Bashkir, Buriat, Chechen, Chuvash, Erzya, Ingush, Kabardian–Circassian, Kalmyk, Karachai–Balkar, Karelian, Khakass, Komi–Peryak and Komi–Zyrian, High and Low Mari, Moksha, Osetin, Russian, Tatar, Tuvin, Udmurt, and Yakut. In Daghestan, the functioning written languages include: Avar, Dargwa, Kumyk, Lak, Lezgi, Nogai, and Tabassaran; their status is guaranteed by the Constitution of Daghestan.

The current language policies are in some ways reminiscent of the early years of the Soviet Union, when citizens were guaranteed education in their native languages. Yet there are significant differences: whereas in the 1920s and 1930s language policy was dictated by a centralized government aiming at unified results, now, in the aftermath of the USSR, these policies are implemented more by local-level governments, within individual republics, with variation from region to region.

In Soviet times, the use of Russian was successfully promoted in Ukraine (*see* **Ukraine: Language Situation**), Belorussia, and Moldova (*see* **Moldova: Language Situation**), reinforced through the migration of ethnic Russians into each. By the time of independence, their citizenry showed preferences for a bilingual state, with some level of official status for Russian, if not on a par with the titular language.

The Baltic States

The Baltic States (Estonia, Latvia, and Lithuania) were the first Soviet Union Republics to enact their own language laws in the late Soviet period. Although the territory of the three Baltic states – Estonia, Latvia, and Lithuania – had been historically the subject of much dispute, the Baltics officially became part of the Soviet Union only in the post World War II period. Their late entry to the USSR meant that the titular languages were already fully established as languages of government and education, and were relatively well-positioned to supplant Russian at the end of the Soviet period. That said, Russian immigration into Estonia and Latvia during the Soviet era left both regions with large numbers of monolingual Russian speakers, something which has slowed transition to exclusive use of the titular languages in government and education.

Both Estonia (*see* **Estonia: Language Situation**) and Latvia (*see* **Latvia: Language Situation**) have instituted specific legislation that bolstered the titular languages and made Russian monolingualism difficult to maintain. In Estonia, a language law which took effect on April 1, 1995, required all people in Estonia to communicate with officials in Estonian only, and to pay for interpreter services if needed. Similarly in Latvia, the 1990s saw legislation which cut support to Russian-language schools and required the use of Latvian only. In Lithuania (*see* **Lithuania: Language Situation**), ethnic Russians constitute a considerably smaller percentage of the population which is, as a whole, more ethnically homogeneous. Although the situation has been calmer there, a move toward Lithuanian monolingualism is also pronounced. One area of direct relation to language use is citizenship. In Lithuania, all residents were automatically granted Lithuanian citizenship. In Estonia and Latvia, however, where nontitular minorities constitute a larger part of the population, only those people who held citizenship in 1940, and their descendants, automatically became citizens. Others are required to go through a naturalization process which

requires, among other things, proficiency in Estonian or Latvian.

Armenia and Georgia

Armenia (Armenia) and Georgia (Georgia) are somewhat distinct from the other successor states and from one another. Both Armenian and Georgian are written in their own distinctive alphabets, and these two languages are the sole examples of languages that were not required to change to the Cyrillic alphabet in the late 1930s. They have strong literary traditions and used their titular languages in education and administration continuously throughout the Soviet period, so that little transition was needed in the immediate post-Soviet years. (Georgian was in fact granted official status in the 1924 Constitution and declared a state language in 1936.) Armenia is more ethnolinguistically homogenous than Georgia and faces few challenges to the Armenian language. In Georgia, linguistic tension is found more between competing local language groups than between Georgian and Russian.

Central Asia and the Turkic States

In 1989–1990, language laws were passed in each of the five Republics of Central Asia, beginning with what is now Tajikistan (*see* **Tajikistan: Language Situation**), with Kazakhstan (*see* **Kazakhstan: Language Situation**), Kyrgyzstan (*see* **Kyrgyzstan: Language Situation**), Uzbekistan (*see* **Uzbekistan: Language Situation**), and Turkmenistan (*see* **Turkmenistan: Language Situation**) all following within a year as already mentioned. These laws differ in particulars but are all similar in instating the titular language as the official state language of each region and attempting to regulate the choice and use of language in public and official settings.

Four of the five Central Asian States (all but Tajikistan) have as their titular language a Turkic language, as does Azerbaijan. The languages of these five Turkic states were written using the Arabic script prior to their incorporation into the Soviet Union. Tajik, a southwestern Iranian (Indo-European) language, was written using the Arabic-based Persian script. In the 1920s, the Soviet government mandated the change from Arabic-based orthographies to Roman-based; this continued until the late 1930s when alphabets throughout the country (except in Armenia and Georgia) were switched to Cyrillic, with the change officially legislated in 1940. With the downfall of the USSR, these States began debating another change of alphabet. Advocates of a pan-Turkic identity promoted a Latin-based orthography to unify Turkic speakers, while others advocated a return to the Arabic-based script. (This was seen as having the potential to unite Azeri speakers in Azerbaijan with those in Iran, who have continuously used the Arabic script.) Three Turkic-speaking States ultimately adopted legislation to switch to Latin-based orthographies: Azerbaijan in 1991; and Turkmenistan and Uzbekistan in 1993, with subsequent laws specifying the rate of change and the exact nature of the alphabet to be used. The transition to the Roman alphabet has been gradual, and it is used together with Cyrillic. In both Kazakhstan and Kyrgyzstan, perhaps due to consideration of the Russian speakers in these regions, no official decision had been made to change the alphabet as of 2004. In Tajikistan, despite a 1989 decree that measures should be undertaken to facilitate learning the Arabic-based Persian script, a revised version of Cyrillic was adopted in 1998.

See also: Education in a Former Colonial Language; Education in a Multilingual Society; Estonia: Language Situation; Kazakhstan: Language Situation; Kirghiz; Kyrgyzstan: Language Situation; Language Politics; Latvia: Language Situation; Linguistic Decolonization; Linguistic Rights; Lithuania: Language Situation; Moldova: Language Situation; Russian Federation: Language Situation; Tajikistan: Language Situation; Turkmenistan: Language Situation; Ukraine: Language Situation; Uzbekistan: Language Situation.

Bibliography

Grenoble L A (2003). *Language policy in the Soviet Union.* Dordrecht: Kluwer.

Kirkwood M (ed.) (1990). *Language planning in the Soviet Union.* New York: St Martin's Press.

Landau J M & Kellner-Heinkele B (2001). *Politics of language in the ex-Soviet Muslim states: Azerbayjan, Uzbekistan, Kazakhstan, Kyrgyzstan, Turkmenistan, and Tajikistan.* Ann Arbor: University of Michigan Press.

Schlyter B N (2001). 'Language policies in present-day Central Asia.' *MOST Journal on Multicultural Societies 3.* http://www.unesco.org.

Wright S (ed.) (2000). *Language policy and language issues in the successor states of the former USSR.* Buffalo: Multilingual Matters.

Language Education, Computer-Assisted

G Davies, Thames Valley University, London, UK

A Definition of CALL

Computers are widely used in the teaching and learning of Modern Foreign Languages (MFL) and English for Speakers of Other Languages (ESOL) in secondary and higher education. The use of computers in teaching and learning languages encompasses many different applications. The applications tend to fall into two distinct types: (a) those that involve the use of generic software tools such as word processors, presentation software, e-mail packages, and Web browsers, and (b) those designed specifically to promote language learning. The latter type falls into the category of Computer Assisted Language Learning (CALL), which is the main focus of this article.

CALL is a term that came into favor in the early 1980s. Levy (1997: 1) provides the following succinct definition of CALL:

> Computer Assisted Language Learning (CALL) may be defined as "the search for and study of applications of the computer in language teaching and learning."

This definition is a very broad one. A more precise way of describing CALL is to say that it is an approach to language teaching and learning in which computer technology is used as an aid to the presentation, reinforcement, and assessment of material to be learned, usually including a substantial interactive element.

A Brief History of CALL

CALL's origins can be traced back to early experiments in the 1960s. Up until the late 1970s, CALL projects were confined mainly to universities, where computer programs were developed on mainframe computers. The PLATO project, initiated at the University of Illinois in 1960, is an important landmark in the early development of CALL (Marty, 1981).

Early CALL (see 'Early CALL' below) favored an approach that drew heavily on practices associated with programmed learning. This aim was reflected in the term Computer Assisted Language Instruction (CALI), which originated in the United States. The term CALI is still used, but it is often associated with a behavioristic approach to language learning that is currently out of favor.

The advent of microcomputers in the late 1970s brought computing within the range of a wider audience, resulting in a boom in the development of CALL programs and a flurry of publications in the early 1980s (Davies and Higgins, 1982, 1985; Kenning and Kenning, 1984; Last, 1984; Ahmad *et al.*, 1985). Most of the CALL programs that were produced in the early 1980s consisted of a series of drills, multiple-choice exercises, and Cloze exercises, focusing on grammar and vocabulary. These kinds of programs were out of tune with orthodox language teaching methodology, which by this time had embraced the communicative approach. There was initially a lack of imagination on the part of CALL developers, a situation that was rectified to a considerable extent by the publication of an influential seminal work by Higgins and Johns (1984), which contains numerous examples of nondrill-based approaches to CALL.

Throughout the 1980s, CALL widened its scope. An alternative term to CALL emerged in the early 1990s, namely Technology Enhanced Language Learning (TELL), which was felt to provide a more accurate description of the activities that fall broadly within the range of CALL.

A major landmark in the history of CALL was the advent of the World Wide Web, which has become the central focus of an increasing number of CALL practitioners and is discussed in more detail in the section 'Web-based CALL' below.

For further information on the history of CALL, see the History of CALL website.

CALL Typology

Various attempts have been made to produce a definitive CALL typology. Davies and Higgins (1982, 1985) distinguish between (a) programs that focus on traditional exercise types such as multiple-choice, gap-filling, free-format, reordering, and Cloze; and (b) programs such as simulations and adventures, action mazes, text manipulation, and exploratory programs. Jones and Fortescue (1987) categorize CALL programs according to those focusing on (a) grammar or vocabulary; (b) the four essential language skills of reading, writing, listening, speaking; with (c) separate categories for adventures, exploratory programs, and authoring programs.

Hardisty and Windeatt (1989) distinguish four broad categories of CALL:

1. School: traditional exercise types
2. Office: generic programs such as word processors and communications software
3. Library: concordancing programs
4. Home: games such as simulations and adventures

Warschauer (1996) looks at CALL in terms of its historical development:

1. Behavioristic CALL: The computer as tutor, serving mainly as a vehicle for delivering instructional materials to the learner.
2. Communicative CALL: With the advent of the microcomputer in the late 1970s, CALL continues to be used for skill practice, but in a non-drill format, with a greater degree of student choice, control, and interaction. This phase also includes (a) using the computer as a stimulus for discussion, writing, or critical thinking; and (b) using the computer as a tool or workhorse – examples include word processors, spelling, and grammar checkers, and concordancing programs.
3. Integrative CALL: This phase is marked by two important technological developments:
   - Multimedia personal computers (MPCs), which enabled reading, writing, speaking, and listening to be combined in a single activity, with the learner exercising a high degree of control over the path through the learning materials.
   - The Internet, which offered new opportunities for computer-mediated communication (CMC) between learners and teachers, and a wide range of activities centered on the World Wide Web.

We shall distinguish the following categories:

- Early CALL
- Communicative CALL
- Multimedia CALL
- Web-based CALL
- CALL authoring programs
- Intelligent CALL (ICALL)
- Computer Aided Assessment (CAA)
- Whole-class teaching and CALL

Early CALL

Early CALL programs – many of which fell into Warschauer's behavioristic category (Warschauer, 1996) – simply presented a series of stimuli to which the learner had to respond. In CALL programs developed for mainframe computers and early microcomputers in the 1960s and 1970s, the stimulus was usually in the form of text displayed in monochrome on a computer screen, and the only way in which the learner could respond was by entering an answer at the keyboard, either by entering a letter or number as a response in a multiple-choice exercise or by typing the answer in full as a response in a free-format exercise. Such programs consisted mainly of a sequence of repetitive drills.

The advent of the microcomputer in the late 1970s gave rise to more imaginative programs, in which text was presented in more interesting ways. Color was used to highlight grammatical features, for example, gender in French and case endings in German. Movement was used to illustrate points of syntax such as position of adjectives in French and subordinate clause word order in German. Discrete error analysis and feedback were a feature of CALL in this period, and the more sophisticated programs attempted to analyze the learner's response, pinpoint errors, and branch to help and remedial activities. A typical example of this approach is the CLEF package for learners of French, which was developed in the late 1970s and early 1980s by a consortium of Canadian universities, and is still widely used in an updated version (Figure 1).

The approach to error analysis adopted by the authors of CLEF – who were predominantly language teachers – was to anticipate common errors and build in appropriately tailored feedback. An alternative approach is the use of Artificial Intelligence (AI) techniques to parse the learner's response, which is a feature of Intelligent CALL (ICALL) – see the section 'Intelligent CALL' (ICALL) below.

Communicative Call

Within the language-teaching profession, there has always been a degree of controversy about the teacher-centered, drill-based approach to CALL. The survey conducted by Levy among CALL developers in the early 1990s indicated that the communicative approach was "the current preferred philosophy of language teaching and learning," but there was also evidence of a high degree of eclecticism, with the majority of teachers adopting two or more different approaches, including the use of drills and teaching formal grammar (Levy, 1997: 122ff.).

Prior to Levy's survey, it was evident as early as the mid-1980s that there was a discernible move away from a drill-based approach to CALL and towards a learner-centered, task-based, or explorative approach. Higgins and Johns (1984: 8–9) were among the first CALL practitioners to illustrate how the computer might adopt different roles other than that of a provider of drills, for example, the computer as an informant, with the learner asking the questions and the computer providing the answers. Jones C. (1986) questioned the role of the computer as a rival to the teacher and stressed the importance of integrating the computer into normal classroom activities, urging the teacher to apply the same methodological expertise and imagination to computer programs as they would to any other classroom aid. Jones G. (1986) and Piper (1986) illustrated how the computer might be used as a stimulus for conversation. In other words, there was a clear break with the top-down programmed learning approach to CALL that characterized its early days.

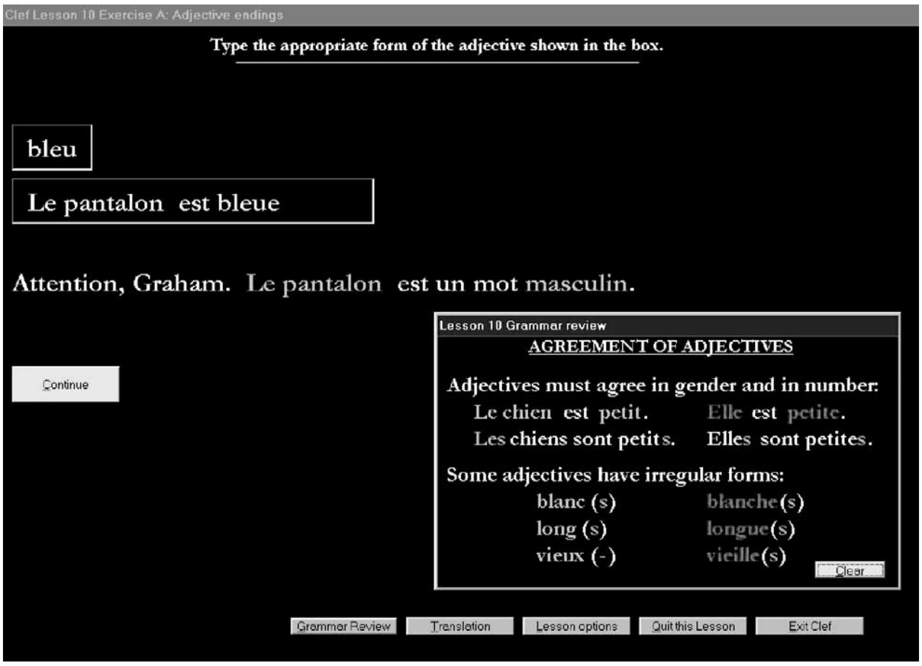


Figure 1 *CLEF* Screenshot. The screenshot is taken from the *CLEF* French grammar CD-ROM (University of Guelph/Camsoft). It shows a fill-in-the-blank exercise on the agreement of adjectives in French, where the learner has entered the incorrect answer 'bleue' (blue), the feminine form of the adjective, instead of the masculine form that is required to agree with the masculine noun 'pantalon' (trousers/pants). The discrete feedback, 'Attention, Graham. Le pantalon est un mot masculin,' points the learner in the right direction rather than giving the correct answer immediately, indicating that the learner needs to make the adjective agree with a masculine noun. The box in the lower right hand corner of the screen is a reminder of the rule about agreement and the forms that adjectives can take. This reminder can be called up by the learner at any time in the course of working on the exercises. Screenshot courtesy University of Guelph/Camsoft.

The explorative approach to CALL is also characterized by the use of concordance programs (concordancers) in the languages classroom. Concordance programs, which were originally used as literary or linguistic research tools, found a new life as tools for the language teacher and student, embodying an approach described by Johns (1991) as Data-Driven Learning (DDL). DDL was the only new approach to CALL that was described in Levy's survey as "a direct result of the attributes of the computer" (1997: 123). See also Tribble and Jones (1990), Johns and King (1991), and Rézeau (2001).

Multimedia CALL

As indicated above in the section Early CALL, computers were initially very limited in terms of the kind of stimulus that they could present to the learner. Most early computers could only offer plain text or simple line drawings. The multimedia personal computer (MPC) was a breakthrough insofar as the stimulus could now consist of any combination of text, still images, sound, and motion video. Ways in which the learner could respond to the stimulus reflected these new developments. Instead of just typing at the keyboard, the learner could point-and-click and drag-and-drop with the mouse, or speak into a microphone. Feedback also reflected the new developments in multimedia.

The earliest manifestations of multimedia CALL were made technologically possible by linking a computer with a 12-inch videodisc player. This enabled text, high-quality photographic images, and sound and video recordings to be presented in a variety of imaginative combinations. The result was the development of interactive videodiscs such as *Montevidisco* (Schneider and Bennion, 1984), *Expodisc* (Davies, 1991) and *A la rencontre de Philippe* (Fuerstenberg, 1993), all of which were designed as simulations in which the learner played a key role.

The techniques learned in the 1980s by the developers of interactive videodiscs were adapted for the integrated MPC, which incorporated a CD-ROM drive and was in widespread use by the early 1990s. The MPC is now the standard form of personal computer. CD-ROMs were used in the 1980s initially to store large quantities of text and later to store sound, still images, and motion video. By the mid-1990s, a wide range of multimedia CD-ROMs for language learners was available. Most CALL programs under development today fall into the category of multimedia CALL.



Figure 2 *World talk* Screenshot. The screenshot is taken from EuroTalk's *World talk* CD-ROM for beginners in French. It is essentially a timed multiple-choice exercise in which three alternative answers to the question below the photograph are presented, taking the form of a TV game show in which the learner has to select the correct answer before his/her opponent 'Matthieu' makes his selection. As well as appearing as text on the computer screen, the question and possible answers are read out by a native-speaker quizmaster, so that the learner can hear the French pronounced as well as reading it from the screen. 'Matthieu' is a video recording and can be seen here reaching for the selection button. Screenshot courtesy of EuroTalk, with permission.

A typical multimedia CALL program includes the presentation of sound and video, usually offering the learner the opportunity to record and playback his/her own voice. An example of a multimedia CALL of this type is EuroTalk's *World talk* program (Figure 2).

The quality of video recordings offered by CD-ROM technology has been slow to catch up with that offered by the earlier interactive videodiscs. The Digital Video Disc or Digital Versatile Disc (DVD), which is now becoming established as a storage medium for multimedia CALL programs, offers much higher quality video recordings, for example EuroTalk's *Movie talk* DVD-ROM series, which is based on authentic TV broadcasts and offers the learner to opportunity to participate in role-plays enhanced by video (Figure 3).

Surprisingly, few imaginative simulations have been developed for the MPC. The series of CD-ROMs entitled *Who is Oscar Lake?* is a notable exception. In this series, the learner is identified as the chief suspect in a diamond robbery in a foreign country and has to argue his/her innocence and determine who the real suspect is. Other multimedia programs make use of Automatic Speech Recognition (ASR) software to diagnose learners' errors, for example Auralog's *Tell me more pro* series (Figure 4).

One of the advantages of multimedia CALL is that it has helped language centers in schools and universities to offer courses that normally would not run due to staff shortages or poor take-up by students. The Critical Language Series, produced by the University of Arizona, offers a number of courses on CD-ROM in languages that are not normally taught as part of the educational curriculum (Figure 5). Similarly, EuroTalk's *Talk now* and *World talk* series offer a wide variety of languages.

Web-based CALL

In 1992, the World Wide Web was launched, reaching the general public in 1993. The Web offers enormous potential and is playing an increasingly important role in language teaching and learning. But, the Web still has some way to go before it catches up with the interactivity and speed of access offered by CD-ROMs or DVD-ROMs, especially when accessing sound and video files. There is no doubt, however, that the Web is a remarkable source of information and has opened up a new range of tasks for learners of foreign languages, e.g., Web quests,



Figure 3 *Movie talk* Screenshot. The screenshot is taken from the Spanish version of EuroTalk's *Movie talk* DVD-ROM for intermediate/advanced learners of Spanish. The DVD-ROM includes a complete episode, *Querido maestro*, from a Spanish TV detective series. The video recording is exploited in different ways, offering sophisticated exercises in which the learner participates in role-plays as well as simple vocabulary exercises like the one illustrated here, in which still images from the recording are presented and the learner has to match the word in the box with one of the six images. As in *World talk* (above), that the learner can hear the French pronounced by a native speaker as well as reading it from the screen. Screenshot courtesy of EuroTalk, with permission.

Web concordancing, and collaborative writing (Felix, 2001).

With the advent of broadband, however, many of the problems associated with the Web are being overcome, but an educational institution needs an extremely fast connection to enable multiple users to enjoy the media-rich language learning materials that are currently available. For this reason, Felix (2001: 190) advises adopting hybrid approaches to CALL, integrating CD-ROMs and the Web and running audio conferencing and video conferencing in conjunction with Web activities.

Web-based multimedia is a growing area of development area but has not yet supplanted CD-ROM or DVD technology. Many websites containing CALL materials offer more in terms of presentation rather than interaction, for example, it is not easy to record and playback one's own voice in a Web environment – a feature of CD-ROM-based multimedia CALL since the late 1980s. Nevertheless, some attractively designed multimedia websites have been produced, for example the BBC Languages website, which offers a comprehensive range of introductory and intermediate courses (Figure 6).

For further examples of what can be done in a Web environment, see LeLoup and Ponterio (2003).

CALL Authoring Programs

Generic authoring packages such as Macromedia's *Director* have been used to create a number of multimedia CALL programs. In the hands of an experienced user, they enable the speedy development of CALL materials, but they are probably too complex for most language teachers and are best suited to the template approach to authoring, as described by Gimeno and Davies (2001). Authoring programs designed specifically for CALL, however, offer a do-it-yourself approach that is within the capabilities of the average language teacher who has no knowledge of computer programming. Typical examples are authoring packages that automatically generate a set of preset activities for the learner, for example, Camsoft's *Fun with texts* and Wida Software's *The authoring suite* (Figure 7). Web authoring packages of a similar type – also offering a range of pre-set activities – are also in widespread use, for example, Hot Potatoes. For further information on CALL authoring programs, see Bangs (2001) and Bickerton *et al.* (2001).

Intelligent CALL (ICALL)

The extent to which the computer is capable of analyzing learners' errors has been a matter of



Figure 4 *Tell me more pro* Screenshot. The screenshot is taken from Auralog's *Tell me more* CD-ROM, a package that incorporates speech-recognition software. It shows a pronunciation exercise in which the learner attempts to repeat the sentence presented in the box at the top of the screen. A native speaker pronounces the sentence and a voice-print of the native speaker model is displayed in the upper of the two voice-print boxes immediately below. The learner's attempt to repeat the sentence is played back and at the same time the learner's voice-print is displayed in the lower box. The learner can attempt the sentence as many times as he/she likes until a close match is achieved. The learner's performance is also displayed in a score box at the lower right hand part of the screen. Screenshot courtesy of Auralog, with permission.

controversy since CALL began. The controversy hinges on those who favor the use of Artificial Intelligence (AI) techniques to develop Intelligent CALL (ICALL) programs, such as Underwood (1989) and Matthews (1992), and its detractors such as Last (1989). The issue of error analysis looms large in this controversy, focusing on the possibilities and limitations of computers programs in diagnosing learners' errors and appropriate feedback (Heift and Schulze, 2003).

Practitioners who come into CALL via the disciplines of computational linguistics, for example, Natural Language Processing (NLP) and Human Language Technologies (HLT), tend to be more optimistic about the potential of these disciplines than those who come into CALL via language teaching. It is clear, however, that recent advances in these areas, such as the development of parsers and speech technology software, are beginning to make a significant impact on CALL (Schulze, 2001; Schulze and Gupta, 2001).

Computer Aided Assessment and Language Learning

Computer Aided Assessment (CAA) covers a range of assessment procedures and is a rapidly developing subset of CALL. CAA refers to any instance in which some aspect of computer technology is used as part of the assessment process. CAA may include tests that take a similar form to CALL exercises, for example, tests in multiple-choice and gap-filling format, but it may also extend to using computers for onscreen marking of students' word-processed writing and using a spreadsheet or database to keep a record of students' grades.

CAA has its limitations, however. While it is particularly useful in testing knowledge of basic grammar and vocabulary and, to a limited extent, reading skills and listening skills, it has not yet had a significant impact on the assessment of speaking skills.

One of the most ambitious CAA projects undertaken in recent years is DIALANG. DIALANG is a major European Union project aimed at providing effective diagnosis of language competence in 14 EU languages. It uses online tests, including placement and self-assessment tests, as key tools in this process. DIALANG is for use by the general public and can be downloaded from the DIALANG website. Tests are available in Listening, Writing, Reading, Structures, and Vocabulary for 14 different languages:



Figure 5 *Critical language series* Screenshot. The screenshot is taken from the opening screen of the Mandarin Chinese CD-ROM that forms part of the *Critical language series* produced by the University of Arizona. It shows the titles of the first 10 lessons of the program. All the programs in the series follow the same pattern, making use of audio and video for presentation, followed by five types of exercises: multiple choice, fill-in-the-blank, audio flashcard, pronunciation, and listening dictation. Learners can record and play back their own voices and compare their pronunciation with that of the native speaker. Screenshot courtesy of University of Arizona, with permission.

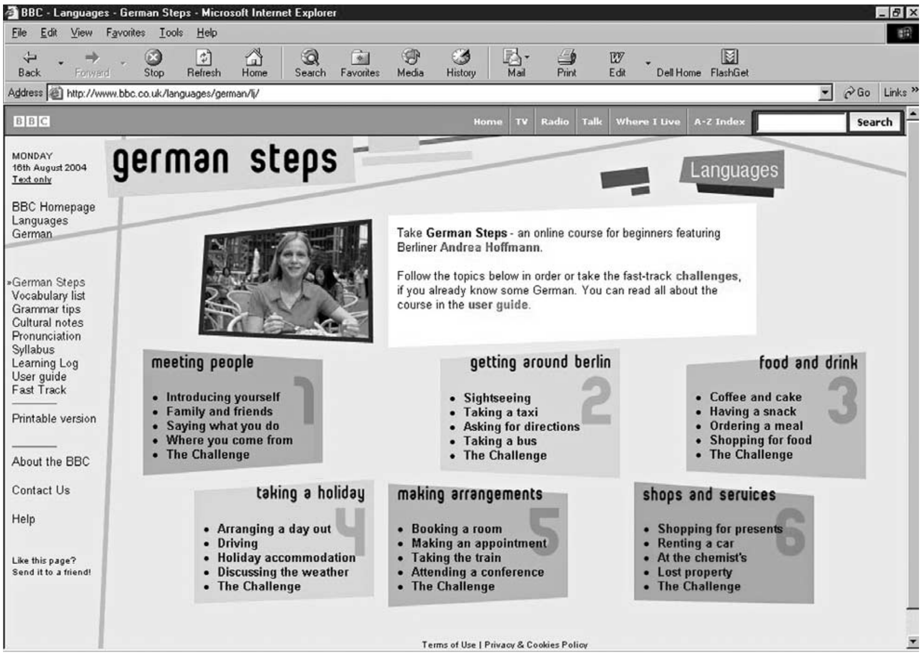


Figure 6 BBC *German steps* Screenshot. The screenshot is taken from the opening screen of the introductory *German steps* course, one of a series produced by the BBC. The screen shows the contents of the course, which consists of materials centered on typical situations that an adult learner might be confronted with on a visit to a German-speaking country. The other programs in the series follow a similar pattern. Screenshot courtesy of BBC, with permission.

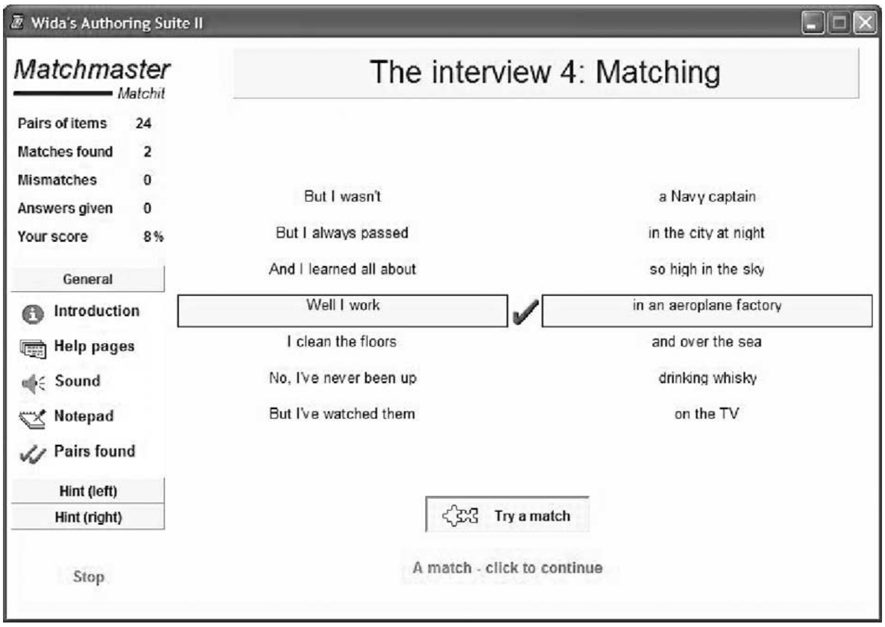


Figure 7 *The authoring suite* Screenshot. The screenshot is taken from Wida Software's *Matchmaster* program, which forms part of a larger package known as *The authoring suite*. This is a matching exercise for learners of English as a Foreign Language, in which the task is to match the beginnings and ends of sentences. *The authoring suite* enables teachers with no programming knowledge to create their own materials. Screenshot courtesy of Wida Software, with permission.

Danish, Dutch, English, Finnish, French, German, Greek, Icelandic, Irish, Italian, Norwegian, Portuguese, Spanish, and Swedish. A key feature of the DIALANG project is that its tests are closely linked to the six levels of the Council of Europe's Common European Framework of Reference for Languages.

Whole-Class Teaching and CALL

When the first computers were introduced into schools, it was usual for a teacher to bring a single computer and a large TV monitor into the classroom and teach the whole class, using the TV monitor as the focus of attention. The teacher or a student would operate the keyboard, and the class would be asked to respond to what appeared on screen. The teacher might use the computer, for example, as a stimulus for eliciting oral responses from the class. This approach worked very well with a variety of programs.

Whole-class teaching went out of favor as computers became cheaper, but it is now undergoing a revival thanks to the advent of the interactive whiteboard, a touch-sensitive projection screen that allows the teacher to control a computer by touching the whiteboard rather than using a keyboard or mouse, although these can still be used. Interactive whiteboards in combination with loudspeakers offer a wide range of activities in the languages classroom, and CALL software designed both for use in computer labs and with interactive whiteboards for whole-class teaching is now widely available, for example mdlsoft's *TaskMagic* package (Figure 8).

The Future of CALL

So, where is CALL heading? Undoubtedly, there will be an expansion of online learning, but it is more likely to supplement conventional modes of learning rather than replacing them. Language learners in particular cannot acquire certain skills, for example, conversational skills, without face-to-face contact with an experienced teacher, but new software tools facilitate synchronous and asynchronous oral computer-mediated communication and are already being used in distance-learning CALL environments.

An area of research and development currently known as Human Language Technologies (HLT) is likely to make an increasing impact on CALL. Schulze and Gupta (2001) describe the main areas of HLT that have already had an influence on CALL and which are likely to have an influence in the not-too-distant future. These areas include Natural Language Processing, Machine Translation, corpus linguistics, and speech technology. All these were once regarded as fringe areas of CALL, but they are now attracting increasing attention.

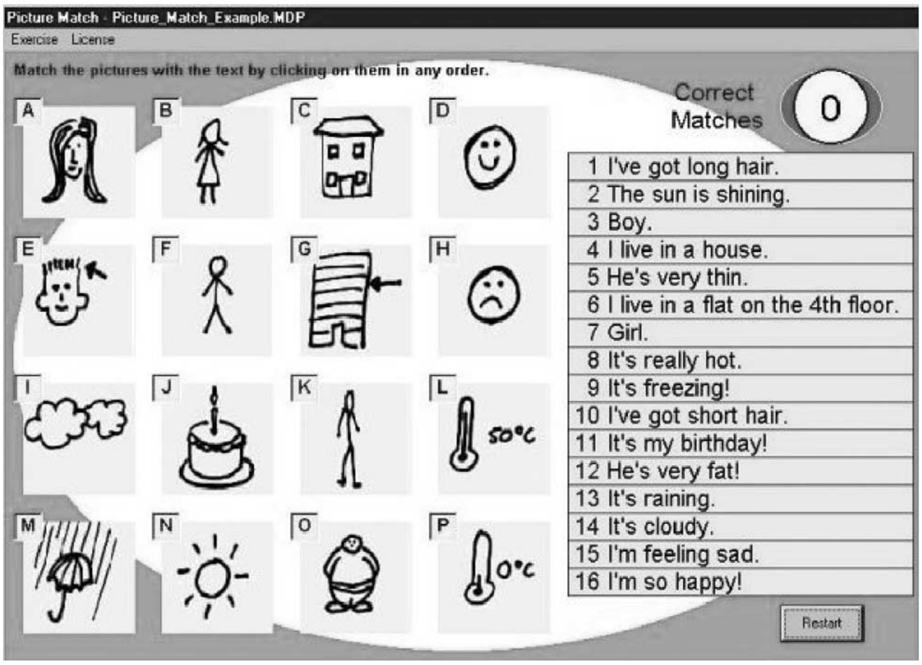


Figure 8 *TaskMagic* Screenshot. The screenshot is taken from mdlsoft's *TaskMagic* package. The aim of this exercise is to match the text in the box on the right of the screen with the pictures. *TaskMagic* is an authoring program that enables teachers with no programming knowledge to create their own materials, and it also lends itself to the creation of materials that are particularly appropriate for whole-class teaching using a computer and a projector or a computer and an interactive whiteboard. Screenshot courtesy of mdlsoft, with permission.

Professional CALL Associations

The rapid growth in the use of computers in language learning and teaching in the 1980s led to the foundation of the two leading professional associations for CALL: CALICO (USA) in 1982 and EUROCALL (Europe) in 1986, both of which continue to thrive. The longer-established professional association, IALLT (founded 1965), focused initially on language laboratories, but later widened its scope to embrace language learning technology in general. An increasing number of professional associations devoted to CALL are emerging worldwide. The established associations are grouped together under WorldCALL, which is in the process of setting itself up as an umbrella association of associations. WorldCALL held its first conference at the University of Melbourne, Australia, in 1998, and the second WorldCALL conference took place in Banff, Canada, 2003.

See also: Behaviorism: Varieties; Communicative Competence; Computational Linguistics: History; Corpus Linguistics; Human Language Technology; Interlanguage; Language Education: Teacher Preparation; Language in Computer-Mediated Communication; Machine Translation: Overview; Natural Language Processing: Overview; Parsing and Grammar Description, Corpus-Based; Teaching Technologies: Second Language.

Bibliography

Ahmad K, Corbett G, Rogers M & Sussex R (1985). *Computers, language learning and language teaching*. Cambridge: Cambridge University Press.

Bangs P (2001). 'Introduction to CALL authoring programs.' Module 2.5 in Davies G D (ed.) *Information and Communications Technology for Language Teachers (ICT4LT)*. Slough: Thames Valley University. 164–183.

Bickerton D, Stenton T & Temmermann M (2001). 'Criteria for the evaluation of authoring tools in language education.' In Chambers A & Davies G D (eds.) *ICT and language learning: a European perspective*. Lisse: Swets & Zeitlinger. 53–66.

Cameron K (ed.) (1998). *Multimedia CALL: theory and practice*. Exeter: Elm Bank Publications.

Council of Europe (2001). *Common European framework of reference: learning, teaching, assessment*. Cambridge: Cambridge University Press.

Davies G D (1991). 'Expodisc – an interactive videodisc package for learners of Spanish.' In Savolainen H & Telenius J (eds.) *EUROCALL 91: Proceedings*. Helsinki: Helsinki School of Economics. 133–139.

Davies G D & Higgins J J (1982). *Computers, language and language learning*. London: CILT.

Davies G D & Higgins J J (1985). *Using computers in language learning: a teacher's guide*. London: CILT.

EUROCALL (2003). *CALL Bibliography*. Available only on-line at http://www.eurocall-languages.org/resources/bibliography.

Felix U (2001). *Beyond Babel: language learning online.* Melbourne: Language Australia.
Fotos S (ed.) (1996). *Multimedia language teaching.* Tokyo: Logos International.
Fuerstenberg G (1993). *A la rencontre de Philippe: videodisc, software, teacher's manual and student activities workbook.* New Haven: Yale University Press.
Gimeno A & Davies G D (2001). 'CALL software design and implementation.' Module 3.2 in Davies G D (ed.) *Information and Communications Technology for Language Teachers (ICT4LT).* Slough: Thames Valley University. 39–78.
Hardisty D & Windeatt S (1989). *CALL.* Oxford: Oxford University Press.
Heift T & Schulze M (eds.) (2003). *Error diagnosis and error correction in CALL 20(3).* Special Issue of *CALICO.*
Higgins J J & Johns T (1984). *Computers in language learning.* London: Collins.
Jones C (1986). 'It's not so much the program: more what you do with it: the importance of methodology in CALL.' *System 14(2),* 171–178.
Jones C & Fortescue S (1987). *Using computers in the language classroom.* Harlow: Longman.
Jones G (1986). 'Computer simulations in language teaching – the KINGDOM experiment.' *System 14(2),* 179–186.
Johns T (1991). 'Should you be persuaded – two samples of Data-Driven Learning materials.' In Johns T & King P (eds.) *Classroom concordancing 4.* Special Issue of *ELR Journal.* 1–16.
Johns T & King P (eds.) (1991). *Classroom concordancing 4.* Special Issue of *ELR Journal.*
Kenning M J & Kenning M M (1984). *An introduction to computer assisted language teaching.* Oxford: Oxford University Press.
Last R W (1984). *Language teaching and the microcomputer.* Oxford: Blackwell.
Last R W (1989). *Artificial Intelligence techniques in language learning.* Chichester: Ellis Horwood.
LeLoup J & Ponterio R (2003). 'Interactive and multimedia techniques in on-line language lessons: a sampler.' *Language Learning and Technology 7(3).* Available online only at http://llt.msu.edu/.
Levy M (1997). *CALL: context and conceptualization.* Oxford: Oxford University Press.
Marty F (1981). 'Reflections on the use of computers in second language acquisition.' *System 9(2),* 85–98.
Matthews C (1992). 'Going AI: foundations of ICALL.' *CALL 5(1–2),* 13–31.
Piper A (1986). 'Conversation and the computer: a study of the conversational spin-off generated among learners of English as a foreign language working in groups.' *System 14(2),* 187–198.
Rézeau J (2001). 'Concordances in the classroom: the evidence of the data.' In Chambers A & Davies G D (eds.) *Information and communications technology: a European perspective.* Lisse: Swets & Zeitlinger. 147–166.
Schneider E W & Bennion J L (1984). 'Veni, Vidi, Vici, via videodisc: a simulator for instructional courseware.' In Wyatt D H (ed.) *Computer assisted language instruction.* Oxford: Pergamon. 41–46.
Schulze M (2001). 'Human Language Technologies in Computer Assisted Language Learning.' In Chambers A & Davies G D (eds.) *Information and communications technology: a European perspective.* Lisse: Swets & Zeitlinger. 111–132.
Schulze M & Gupta P (2001). 'Human Language Technologies.' Module 3.5 in Davies G D (ed.) *Information and Communications Technology for Language Teachers (ICT4LT).* Slough: Thames Valley University. 117–156.
Tribble C & Jones G (1990). *Concordances in the classroom.* Harlow: Longman.
Underwood J (1989). 'On the edge: Intelligent CALL in the 1990s.' *Computers and the Humanities 23,* 71–84.
Warschauer M (1996). 'Computer assisted language learning: an introduction.' In Fotos S (ed.) *Multimedia language teaching.* Tokyo: Logos International. 3–20.

Relevant Websites

http://www.auralog.com/english.html – Auralog, producers of CALL Software, France.
http://www.bbc.co.uk/languages – BBC Languages, a series of Web pages for learners of foreign languages.
http://www.camsoftpartners.co.uk/clef.htm – CLEF, a software package for learners of French, produced by the University of Guelph and published in the UK by Camsoft.
http://clp.arizona.edu/cls – Critical Language Series, a CD-ROM series produced by the University of Arizona.
http://www.dialang.org – DIALANG, a diagnostic language testing project, initiated with funding from the European Union.
http://www.eurotalk.co.uk – EuroTalk, producers of CALL software, London, UK.
http://www.history-of-call.org – History of CALL, a website containing a wealth of information on the history of CALL.
http://web.uvic.ca/hrd/halfbaked/ – Hot Potatoes, a Web-authoring package produced by Half Baked Software Inc, University of Victoria, Canada.
http://www.ict4lt.org – ICT4LT (ICT for Language Teachers), a large collection of ICT training resources for language teachers.
http://llt.msu.edu – Language Learning Technology Journal (LLTJ), an on-line refereed journal.
http://www.languagepub.com – Language Publications Interactive (LPI), producers of *Who is Oscar Lake?* a language simulation on CD-ROM.
http://www.mdlsoft.co.uk – mdlsoft, producers of the *TaskMagic* package.
http://www.wida.co.uk – Wida Software, producers of *The authoring suite.*
http://www.tell.is.ritsumei.ac.jp/cocllejonline/, Kusatsu, Shiga, Japan: Ritsumeikan University, available only on-line (the

CALL-EJ journal (Japan) merged with the *On-CALL* journal (Australia) and became *CALL-EJ Online* in May 1999).
http://www.taylorandfrancisgroup.com – *CALL Journal*, London: Taylor & Francis, available on-line from Volume 11.
http://llt.msu.edu – *Language Learning Technology*, East Lansing: Michigan State University, available only on-line.
http://www.upv.es/worldcall/ – WorldCALL, an umbrella organization of professional associations for CALL.
http://www.eurocall-languages.org – EUROCALL: a European professional association for CALL; Cambridge University Press publishes the journal *ReCALL* on the association's behalf.
http://www.calico.org – CALICO: a North American professional association for CALL and publisher of the *CALICO Journal*.
http://www.iallt.org – IALLT: a North American professional association and publisher of the *IALLT Journal of Language Learning Technologies*.

Language Education: Correctness and Purism

W Davies, University of Wales Aberystwyth, UK

Definitions: Correctness and Purism

The common assumption that school constitutes one sociolinguistic domain – a domain usually dominated by the standard variety of a language – is too simplistic. School consists of various communicative situations, e.g., playtimes, conversations between teacher and pupil outside the classroom, teacher–teacher communication. On the other hand, language education in the sense of the explicit teaching of language takes place almost exclusively in the classroom. In this context, the point of reference for determining whether linguistic usage is 'correct' or not is usually, but not always, a standard variety. In societies where there is a standardized prestige variety this nearly always has a privileged position in the classroom, i.e., it is the variety which is not only used as the medium of teaching, but is also an object of teaching: in Britain or the United States, for instance, in classes devoted to English, it is the norms of Standard English which are transmitted to pupils, and the texts discussed in these classes are mainly in Standard English. Deviations (real or perceived) from this pattern tend to be seized upon by certain sectors of society and presented in a negative light, for example in the media (cf. Cameron, 1995: 101).

Purism is rooted in the notion that languages or linguistic varieties are entities which are clearly demarcated from each other and consist of closed sets of linguistic items, so that elements that do not 'belong' to that variety and which might 'contaminate' it or make it 'impure' are easily identified and can be avoided. Purism is often, but not always, directed at words or linguistic features of foreign origin (cf. Trask, 1999: 254). It can, however, also be directed against elements which are not foreign, but are nevertheless seen as undesirable; for example, purists may try to remove regional dialectal features from a standard variety or may try to discourage the use of innovations, such as *aggravate* with the meaning 'annoy' (appealing to etymology, they would argue that *aggravate* should only be used in the sense of 'to make worse'). Another target of their activities may be words or constructions from specialist registers when used outside those registers, e.g. in German technical registers, mass nouns such as *Druck* ('pressure') can form plurals, whereas this has traditionally not been the case in non-technical language. A tendency for this construction to spread to general registers has often been the subject of negative comments, e.g., in newspaper glosses. Purism is often associated with the codification, cultivation, and planning of standard languages, but this doesn't mean that purism is never directed at non-standard varieties. In German-speaking Switzerland, for example, where the local dialects enjoy a high level of prestige, but where attitudes towards Swiss Standard German are rather ambivalent, purists have been known to attack standard features borrowed into the local dialects.

Purism is a strategy employed to achieve a variety of ends. Sometimes foreign words are rejected because it is claimed that they are not comprehensible to all and could exclude sections of the population from certain discourses; at other times they are rejected solely because of their foreignness, which may be seen as detrimental to the 'character' of a particular nation as reflected in its language (cf. Townson, 1992: 80–110).

As was said above, purism is very much linked to the cultivation of a standard variety. One of the things that lends prestige to a standard variety is the belief that it can fulfill any function demanded of it by its speakers. If it is perceived as being able to fulfill certain functions only by 'borrowing' elements from other varieties, then this detracts from its status.

CALL-EJ journal (Japan) merged with the *On-CALL* journal (Australia) and became *CALL-EJ Online* in May 1999).
http://www.taylorandfrancisgroup.com – *CALL Journal*, London: Taylor & Francis, available on-line from Volume 11.
http://llt.msu.edu – *Language Learning Technology,* East Lansing: Michigan State University, available only on-line.
http://www.upv.es/worldcall/ – WorldCALL, an umbrella organization of professional associations for CALL.
http://www.eurocall-languages.org – EUROCALL: a European professional association for CALL; Cambridge University Press publishes the journal *ReCALL* on the association's behalf.
http://www.calico.org – CALICO: a North American professional association for CALL and publisher of the *CALICO Journal*.
http://www.iallt.org – IALLT: a North American professional association and publisher of the *IALLT Journal of Language Learning Technologies*.

Language Education: Correctness and Purism

W Davies, University of Wales Aberystwyth, UK

Definitions: Correctness and Purism

The common assumption that school constitutes one sociolinguistic domain – a domain usually dominated by the standard variety of a language – is too simplistic. School consists of various communicative situations, e.g., playtimes, conversations between teacher and pupil outside the classroom, teacher–teacher communication. On the other hand, language education in the sense of the explicit teaching of language takes place almost exclusively in the classroom. In this context, the point of reference for determining whether linguistic usage is 'correct' or not is usually, but not always, a standard variety. In societies where there is a standardized prestige variety this nearly always has a privileged position in the classroom, i.e., it is the variety which is not only used as the medium of teaching, but is also an object of teaching: in Britain or the United States, for instance, in classes devoted to English, it is the norms of Standard English which are transmitted to pupils, and the texts discussed in these classes are mainly in Standard English. Deviations (real or perceived) from this pattern tend to be seized upon by certain sectors of society and presented in a negative light, for example in the media (cf. Cameron, 1995: 101).

Purism is rooted in the notion that languages or linguistic varieties are entities which are clearly demarcated from each other and consist of closed sets of linguistic items, so that elements that do not 'belong' to that variety and which might 'contaminate' it or make it 'impure' are easily identified and can be avoided. Purism is often, but not always, directed at words or linguistic features of foreign origin (cf. Trask, 1999: 254). It can, however, also be directed against elements which are not foreign, but are nevertheless seen as undesirable; for example, purists may try to remove regional dialectal features from a standard variety or may try to discourage the use of innovations, such as *aggravate* with the meaning 'annoy' (appealing to etymology, they would argue that *aggravate* should only be used in the sense of 'to make worse'). Another target of their activities may be words or constructions from specialist registers when used outside those registers, e.g. in German technical registers, mass nouns such as *Druck* ('pressure') can form plurals, whereas this has traditionally not been the case in non-technical language. A tendency for this construction to spread to general registers has often been the subject of negative comments, e.g., in newspaper glosses. Purism is often associated with the codification, cultivation, and planning of standard languages, but this doesn't mean that purism is never directed at non-standard varieties. In German-speaking Switzerland, for example, where the local dialects enjoy a high level of prestige, but where attitudes towards Swiss Standard German are rather ambivalent, purists have been known to attack standard features borrowed into the local dialects.

Purism is a strategy employed to achieve a variety of ends. Sometimes foreign words are rejected because it is claimed that they are not comprehensible to all and could exclude sections of the population from certain discourses; at other times they are rejected solely because of their foreignness, which may be seen as detrimental to the 'character' of a particular nation as reflected in its language (cf. Townson, 1992: 80–110).

As was said above, purism is very much linked to the cultivation of a standard variety. One of the things that lends prestige to a standard variety is the belief that it can fulfill any function demanded of it by its speakers. If it is perceived as being able to fulfill certain functions only by 'borrowing' elements from other varieties, then this detracts from its status.

Another important role ascribed to a standard variety is that it is supraregional and has a wide communicative radius: many nonstandard words or constructions are perceived as being restricted in communicative radius and it is therefore felt that they must not be allowed into the standard variety. In reality, it is often the case that the lack of intelligibility of nonstandard variants is socially constructed and not rooted in any objective linguistic distance between those variants and the standard variety. There are various instances of speakers of very similar varieties claiming that they could not understand each other, with speakers of the more prestigious varieties usually more likely to claim not to understand other varieties than *vice versa* (e.g., Wolff, 1959: 35–39).

As we have already said, language education in the classroom is closely linked with standard varieties, so the link with correctness and purism is clear. Deborah Cameron makes it explicit when she claims – for Britain – that "one of the functions that conservatives allot to grammar teaching specifically is the preservation of a 'pure' and unified standard English variety whose purity is continually threatened just because it is not native to most speakers of English" (Cameron, 1995: 110). That this is not restricted to Britain is clear when Steven Pinker, a Canadian based in the United States, lists English teachers amongst what he calls "the legislators of 'correct English'" (Pinker, 1994: 372), and Rosina Lippi-Green writes, "The primary educational goal in our schools brings together the acquisition of literacy with acceptance and acknowledgement of a Standard U.S. English" (Lippi-Green, 1997: 104).

The Appropriateness Model of Linguistic Variation

Correctness and purism come into play not only in relation to particular individual linguistic features (e.g. *ain't* vs. *haven't*, *less items* vs. *fewer items*) but also in relation to the use of different varieties in different situations or for different speech acts, e.g., for writing an essay or for making an oral presentation. There are people who believe that variation is a problem *per se*, since speakers of nonstandard varieties will always be liable to making interference mistakes when they acquire the standard variety, especially in speech. In the view of these people, the only way to acquire a 'correct' or 'pure' form of the standard (which is, of course, seen as an eminently desirable thing) is to eliminate the nonstandard varieties. In the 1970s, this was argued for Germany and it has often been argued with reference to minority languages in order to promote the learning of the majority language (the demise of the Celtic languages has often been encouraged for this reason). A more common view today, especially amongst academic linguists, is that varieties of a language differ from each other and are therefore not all equally appropriate for the same purposes and situations. This model, known as the appropriateness model, is based on the difference theory of linguistic variation. The difference theory (unlike the deficit theory) assumes the functional equivalence of all varieties of a language, whilst accepting that they are differently evaluated by society.

This model seems to offer teachers a way of responding more positively to the variation they inevitably come across every day since they don't have to pass absolute judgments of correctness or incorrectness on their pupils' usage. They can now say, "You can say that at home, but not in the classroom" rather than saying, "Don't say that, it's wrong." The difference theory and the appropriateness model have certainly had an effect on policy makers in various countries, e.g. in Germany and Britain. The Cox Report (1989) on the teaching of English in primary schools in England and Wales, commissioned by the Conservative government of the time, made the point that Standard English should not be confused with 'proper,' 'good,' or 'correct' English. This reflects the view that other varieties of English may be seen as better or more proper (i.e., more appropriate) in some situations than Standard English; the status of the latter is relativized: "The combination of the status of Standard English as a social dialect and its use for particular purposes ... poses the most sensitive problems for teachers. They have the responsibility to teach its use for those particular purposes, while being sensitive to its impact on personal and social identity" (DES, 1989: 14–15). Whilst not saying that the standard variety of English is better than other varieties, the Cox Report still puts a great deal of emphasis on Standard English (written and spoken), and on the requirement that it be used in a range of situations.

Although the appropriateness model is popular with academic linguists, it does not appeal to all policy makers, especially those of a conservative bent, since it implies that there are no absolute values and standards of correctness. Another tenet of academic linguistics – that all varieties are rule governed, even if those rules are not codified or promoted by institutions such as the media or schools – is highly suspect to some policy makers since it takes away the notion of an external authority and makes it more difficult to apply uniform rules of 'correctness' (cf. Cameron, 1995: 97–98). There is a great deal of resistance to the fact that nonstandard varieties have a 'grammar,' although it is rarely written down. This situation is not helped by the fact that there is a dearth

of systematic descriptions of nonstandard varieties of English. It is not easy, for instance, for a teacher in the United States to find out that, in Black English Vernacular (BEV), invariant *be* is not simply the equivalent of inflected *be* in Standard U.S. English, but has a more specific distribution; e.g., *He be busy right now* does mean the same thing as *He is busy right now*, but Standard U.S. English *He is busy all the time* can only be rendered as *He busy all the time* in BEV (with zero copula, but not with invariant *be*). Of course, the resistance to pluralism in the linguistic sphere is often accompanied by a resistance to pluralism in other fields, and the former is often the butt of particularly fierce criticism because it is seen as not only reflecting but positively encouraging cultural and ideological fragmentation.

Despite some opposition over the years, proponents of the appropriateness model seem to have been quite successful in influencing school curricula. For example, in Germany (where each federal state is responsible for its own curriculum), most curricula refer to the fact that pupils should be taught to use different varieties of German in an 'appropriate' fashion (cf. Davies, 2000a; Davies, 2000b; Davies, 2001). Richard Hudson also points out that the National Curriculum for English in England and Wales accepts the importance of oracy, as well as literacy. Furthermore, references to 'genre' in the curriculum reflect an acceptance that many varieties exist and are 'valid.' (cf. 'Relevent Websites'). In Wales, where Welsh, which is spoken by about 20% of the population, is compulsory in schools up to the age of 16, stress is put on the use of appropriate registers; e.g., different forms are required when creating a script for a drama and when writing an official letter (Robat Powell, Head of National Foundation for Educational Research Welsh Unit, personal correspondence).

While the appropriateness model has been widely accepted by policy makers, calls by academic linguists for a critical approach to the domain distribution of different varieties and the social evaluation underpinning it have not found the same echo. The traditional distribution of standard and nonstandard varieties across public and private domains or formal and informal ones, has, on the whole, been accepted rather than challenged within that model, with the result that the dominant position of the standard variety has hardly been affected. Research in Germany (Davies, 2000b) has shown that neither teachers nor the curricula show any enthusiasm for extending the permitted domains of nonstandard varieties. In theory at least, the privileged status of the standard language in the classroom is not questioned by the practitioners any more than by the theorists.

In Britain and the United States, there have often been perceptions that teachers are trying to replace Standard English with nonstandard varieties, or are not concerned enough about correctness. The British heir apparent, Prince Charles, is one member of the public who worries about the state of language education in British schools: "All the letters sent from my office I have to correct myself, and that is because English is taught so bloody badly [. . .]. The way schools are operating is not right. I do not believe English is being taught properly" (quoted in Cameron, 1995: 94). In California, when the Oakland School Board resolved in 1996 that language education should respect and take into account the vernacular of most of its pupils, namely Ebonics (another name for BEV), there was an extremely hostile reaction from the general public, but this was fueled by the misleading reports in the media that the schools in this district were intending to teach Ebonics and/or use it as the medium of instruction (information about this issue can be seen at the web site of the Linguistic Society of America). In reality, when nonstandard varieties are allowed into the classroom, it is usually in order to create a bridge between the home or the peer group and the classroom. The ultimate aim for most teachers is the expansion of the linguistic range of their pupils through the acquisition of the standard variety. This is reflected in the resolution of the Linguistic Society of America on the Oakland 'Ebonics' issue: "There is evidence from Sweden, the United States, and other countries that speakers of other varieties can be aided in their learning of the standard variety by pedagogical approaches which recognize the legitimacy of the other varieties of a language."

What is a Standard Variety?

So far we have used the term 'standard variety' as though it were so well known that it needed no definition; however, if the curricula use the standard variety as their point of reference then teachers have to know exactly what the standard variety is, and this may not be as self-evident as one would assume, even for a language where a standard prestige variety has been recognized for many centuries, e.g., English in Britain. Some standard varieties are highly codified: for example, the national variety of German used in Germany has a whole range of reference works in which speakers can look up rules for 'correct' use of phonology, orthography, grammar, and lexis. These works are produced by commercial publishers (the best known is Duden-Verlag, based in Mannheim) rather than by a state institution (like, for example, the *Österreichische Wörterbuch* in Austria), but are still perceived by many users as officially binding.

Indeed, up until the spelling reform a few years ago, the *Duden Rechtschreibung* (the orthography dictionary) did have the authority of the education ministries of all the federal states behind it. The various national varieties of English are not codified to this degree and the norms are more in the nature of conventions that have emerged in practice rather than codified rules (cf. Durrell, 1999: 290). In the United States, there are competing perceptions of what constitutes spoken Standard U.S. English and where it is to be found (cf. Hartley and Preston, 1999). Speakers from Michigan and California seem to be especially sure that the varieties they use represent pure, Standard U.S. English, even when there are nonstandard features in these varieties (as judged by the informants themselves) (Fought, 2002; Niedzielski and Preston, 1999). It is sometimes claimed that 'network American' is Standard U.S. English: this is a mainstream accent, based on the leveled dialects of the northern Midwest, with the most marked local features being avoided (Milroy and Milroy, 1999: 150–151).

When it comes to British English there seems to be agreement amongst linguists that it is quite difficult to describe the rules of Standard English unless one restricts oneself to the written realization only. It is accepted that the spoken variety cannot be standardized to the same degree, and that even speakers who avoid the use of locally or regionally marked nonstandard grammar and vocabulary will nearly always still have a regional accent (cf. Milroy and Milroy, 1999: 18–19). It is clear from the preceding that academic linguists conceive of Standard English as being realized in a spoken and written form, although policy makers (or even norm makers) frequently fail to draw such a distinction (cf. Lippi-Green, 1997: 107) and often see differences from writing as deviations from a written norm rather than judging them on their own terms.

A problem for teachers is that the curricula which they are expected to follow often give no clear guidance as to what is to be taught under the guise of Standard English, German, Welsh, etc. For teachers of Welsh in Wales the situation is as follows: there are "no instructions or guidelines concerning the grammatical and syntactic forms to be used, except for a book of terms where there are lists of 'technical' words classed according to subject ... Every type of Welsh is accepted at present" (Derec Stockley, Examinations Officer of the Welsh Joint Education Committee, which sets exams in Welsh as well as other subjects, quoted in Jones, 1998: 282). The lack of explicit guidelines in the curriculum is perhaps not such a problem in a country such as Germany where there is a codex (although the fact that a codex exists does not necessarily mean that the teachers are familiar with its content; cf. Davies, 2000b); in this case it is, however, potentially problematic that the codex almost exclusively contains the rules for a formal register of German (Durrell, 1999: 302–303) and makes few concessions to spoken usages, even those common amongst educated speakers in formal situations (Jäger, 1980: 376).

Even if the curricula of teacher training colleges are to accept that their trainees need to acquire knowledge of the systematic differences between the nonstandard varieties which the majority of their pupils will speak and the standard variety, the information is often simply not available, or is not particularly useful. As we saw above, not all language communities possess full descriptions of their written and spoken standard varieties, but the problem is even more acute for nonstandard varieties, and there is a pressing need for more research in order to have up-to-date and detailed descriptions of nonstandard features at all levels of language. Jenny Cheshire (1982: 53) found that teachers of English who lacked awareness of the systematic differences between Standard English and the nonstandard dialect of their pupils tended to mark inconsistently, thereby confusing the pupils. In Germany, between 1977 and 1981, a range of books presenting contrastive analyses of Standard German and regional nonstandard dialects was published (*Dialekt-Hochsprache kontrastiv*), aimed at teachers to help them to teach Standard German to their pupils. Unfortunately, these were of limited use since the analyses were based on a comparison between formal Standard German and traditional dialect, whereas most of the children, especially in urban areas, spoke an intermediate form at home (cf. Barbour, 1987).

The lack of detailed descriptions of different linguistic varieties is also a problem in the United States, where standardized language tests are often used to help determine the kind of education most appropriate to a child encountering problems at school (Milroy and Milroy, 1999: 143). If the correctness of a child's answer is measured solely against one Standard English form and if every deviation from that form is marked as wrong, regardless of whether it is part of the child's vernacular or not, then the results are going to be badly skewed and unhelpful, and may well lead to underestimation of the child's real level of linguistic skills. An example is the contrast between /f/ and /θ/ (*fin/thin*). Not all dialects of English have minimal pairs based on this distinction, so a child who speaks one of those dialects is more likely to filter out the auditory information s/he needs in order to perceive the contrast than is a child whose dialect makes a phonemic contrast between the two

sounds. This is comparable to a native speaker of Standard British English who has to practice hearing the difference between *wurde* and *würde* in German since his/her dialect doesn't have front rounded vowels. In the United States, speakers of nonstandard varieties score very much lower on language tests than do speakers of the standard variety, suggesting that there is a bias in the test material and that tests do not always allow a clear distinction to be made between children who have underdeveloped language skills and those who may not have mastered the standard variety, but who are perfectly fluent in their vernacular (Milroy and Milroy, 1999: 143).

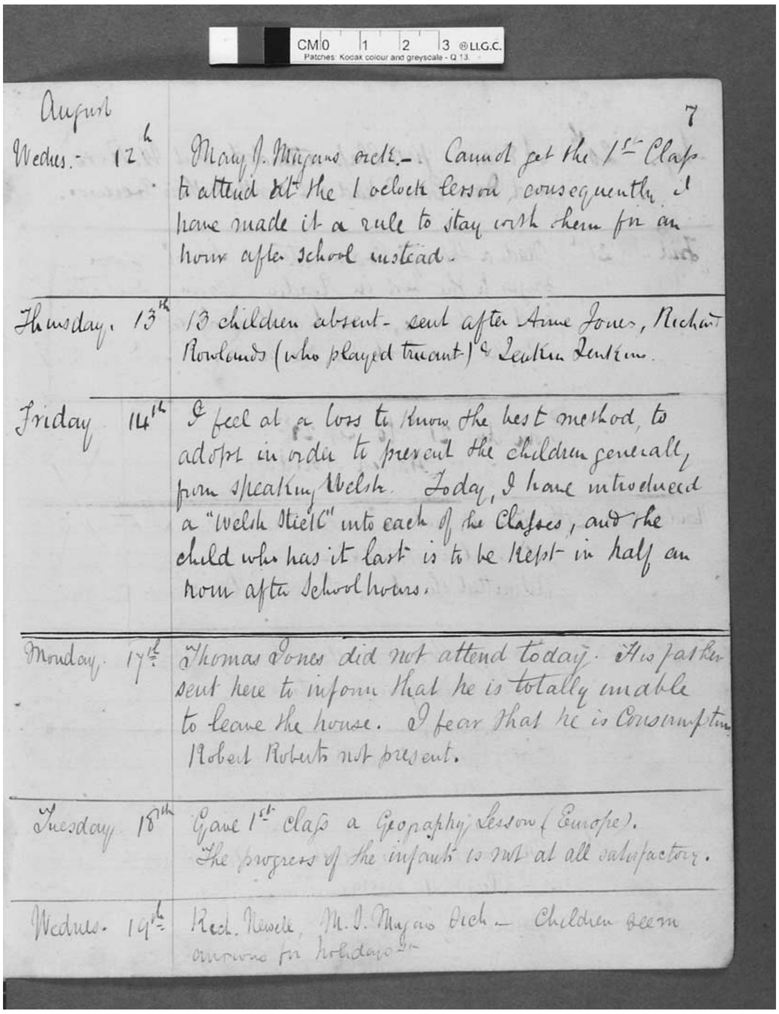
Spoken and Written Language

As we saw above, there is often a tendency to equate the standard variety of a language with the formal written registers of that language. This may make a teacher's life easier to some extent, since the written realization of the standard variety is normally relatively fixed and invariant, and this provides an authoritative norm against which to measure a child's performance. However, if we measure children's oral performances against the norms of written language they are always going to be found wanting, even those who come from homes where a standard-like variety is spoken. Even these children have to learn the rules of written Standard English or French or German, etc.; they cannot simply write as they speak since no variety of speech is exactly the same as written language. Teachers need to be able to differentiate between errors in written work that are due to interference from regional nonstandard dialects and those that are due to interference from informal colloquial varieties. This is not an easy task: not only is it the case, as we have already seen, that spoken Standard English is not easy to define, but, as Cheshire points out, "the grammatical structure of spoken English generally is far from being well understood" (Cheshire, 1999: 129). The syntax of spoken language is very different from that of writing, and constructions that appear to be redundant or ill formed according to the norms of writing play important roles in structuring spoken discourse.

One reason why children need to learn to distinguish between the requirements of speech and writing is that the latter, since it is relatively context free, cannot tolerate as much ambiguity or vagueness as the former. Uniformity is also useful in writing since it helps to ensure that the text will be understood over a wide area and by a wide range of people. The same degree of explicitness is not necessary in speech, although teachers often act as if it were, e.g., by asking children to answer in full sentences, when this is quite unusual in normal interaction. In highly literate societies like the United States, Britain, or Germany there tends to be an assumption that clear and effective communication can only take place in a standard variety, and then in its most formal realization (usually the realization closest to writing). People often find it difficult to accept that clear and effective communication can take place in a nonstandard variety, and it is clear that, in certain situations or with certain interlocutors, communication will not be as effective in a nonstandard variety since the listeners will be so distracted by the nonprestigious form that they may not pay the necessary attention to the content of the message. However, this is related to social value judgments and is not inherently a property of the nonstandard varieties themselves; e.g., not pronouncing orthographic *h* in words like *hat* may be seen as incorrect and judged negatively in certain (influential) sectors of British society, but the same phenomenon in Spanish is part of the standard norm (*haber*, *honor*, etc.).

Multilingualism

Another major issue that is connected with correctness and/or purism is the role of languages other than the dominant or official national or state languages in the classroom. In Britain, for example, about 10% of schoolchildren speak English as an additional language (EAL). Many of these children were born in Britain or have been living here for a long time. The prevailing trend is to teach them in mainstream classrooms. Even those children who have recently moved to Britain are taught in mainstream classrooms, but they also receive some support from specialist EAL teachers. In recent years there has been increasing discussion about the necessity of integrating immigrants to Britain, and the role of language education (i.e., learning English) in this process has been hotly debated, although as yet there is no need to pass a language test before being able to obtain British citizenship. There is some pressure for classrooms to become spaces in which only English may be spoken, since there is a perception that English is not used enough in other, more private domains (see Figure 1). A similar movement has been obvious in the United States for some years: the U.S. English organization wants to outlaw the use of languages other than English in public domains. Superficially, such movements seem to have something in common with schools devoted to the maintenance or revitalization of minority or endangered languages; for example, in Welsh-medium schools in Wales and in French-medium schools where French is the minority language (e.g., in the Canadian province of Ontario)



August 7

Wednes.- 12th Mary J. Mygans sick.- Cannot get the 1st Class to attend till the 1 oclock lesson consequently I have made it a rule to stay with them for an hour after school instead.

Thursday. 13th 13 children absent- sent after Anne Jones, Richard Rowlands (who played truant) & Jenkin Jenkins.

Friday 14th I feel at a loss to know the best method, to adopt in order to prevent the children generally from speaking Welsh. Today, I have introduced a "Welsh Stick" into each of the Classes, and the child who has it last is to be kept in half an hour after School hours.

Monday. 17th Thomas Jones did not attend today. His father sent here to inform that he is totally unable to leave the house. I fear that he is Consumptive. Robert Roberts not present.

Tuesday 18th Gave 1st Class a Geography Lesson (Europe). The progress of the infants is not at all satisfactory.

Wednes. 19th Recd. Newell, M. J. Mygans sick — Children seem anxious for holidays.

Figure 1 This extract from log book of Towyn British School, Merionethshire, Wales 1863–1876, referring to attempts to prevent children from speaking Welsh, shows that such pressure is not new. From Meirionnydd Archives, Gwynedd Archive Service with permission.

great efforts are made to try to ensure that only Welsh or French is spoken at school. Whether the strategy is necessary in the case of English, the dominant language in Britain and the United States, is debatable.

Studies of code-switching in the classroom have thrown a great deal of light on the hierarchical relations between different languages in multilingual societies. In some societies, e.g., Malta, code-switching between English and Maltese is officially encouraged and is common in most classrooms, although the exact distribution of the languages will depend a great deal on the individual teacher (cf. Camilleri Grima, 2001). However, since there are far more teaching materials available in English than in Maltese, the already high status of English in Maltese society is reinforced in the classroom, and Maltese tends to be used for less prestigious activities, such as to create an informal atmosphere or to create warmth (Camilleri Grima, 2001). In Botswana, the official policy is that all education from the last few years of primary schooling onwards is to be in the postcolonial language, English, rather then the indigenous language Setswana (Tswana). Teachers occasionally switch between the two languages, but for pupils the only legitimate language is English (Arthur, 2001). In British classrooms, bilingual teaching assistants are sometimes employed in the early years of primary education to help minority-language students become integrated into school life; however, these assistants are rarely drawn into the main action of the classroom, but left on the margins while the main action takes place monolingually in English (Martin-Jones and Saxena, 2001).

Policy vs. Practice

Official policies as expressed in curricula and similar documents say a great deal about official views of how language education should be realized, but we cannot assume that the practitioners slavishly follow such policies. Research in Germany has shown that, even when the curricula state clearly that Standard German is always to be used as the medium of instruction, and although teachers do not explicitly question the privileged status of Standard German at school, the classroom is nonetheless not exclusively the domain of the standard. Teachers do not always insist that their pupils speak the standard unless they are performing certain tasks, such as giving oral presentations, and they often express an awareness that correcting children too often is likely to discourage them from contributing in class. There is also an awareness that some children might find it difficult to perform certain speech acts, e.g., telling a teacher about a personal problem, in the standard, unless this is their mother tongue (cf. Davies, 2000b). What is clear, however, is that the use of nonstandard varieties in the classroom is usually restricted to rather stereotypical topics such as dialect literature, or to speech events that are marginal to the 'real' business of teaching (personal conversations), or to nonacademic topics, and the dominant status of Standard German is never put in doubt. Similarly, in Malta, as we saw above, although official discourse presents both languages as equal, the reality is that Maltese is used for less prestigious speech acts than English.

See also: Assessment of First Language Proficiency; Bilingualism; Code Switching; Discrimination and Language; Duden, Konrad (1829–1911); Education in a Multilingual Society; Educational Failure; Germany: Language Situation; Identity and Language; Language Attitudes; Language Education Policies in North America; Language Education Policy in Europe; Language Education: Teacher Preparation; Language Policy in Multilingual Educational Contexts; Minorities and Language; Minority Languages: Education; Minority Languages: Oppression; Mother Tongue Education: Nonstandard Language; Mother Tongue Education: Standard Language; Norms and Correctness; Prestige, Overt and Covert; Society and Language: Overview; Standardization; Teaching of Minority Languages; United Kingdom: Language Situation; United States of America: Language Situation; Variation in German; Wales: Language Situation.

Bibliography

Arthur J (2001). 'Codeswitching and collusion: classroom interaction in Botswana primary schools.' In Heller M & Martin-Jones M (eds.). 57–75.

Barbour S (1987). 'Dialects and the teaching of a standard language: some West German work.' *Language in Society 16,* 227–243.

Cameron D (1995). *Verbal hygiene.* London/New York: Routledge.

Camilleri Grima A (2001). 'Language values and identities: codeswitching in secondary classrooms in Malta.' In Heller M & Martin-Jones M (eds.). 213–234.

Cheshire J (1982). 'Dialect features and linguistic conflict in schools.' *Educational Review 34,* 53–67.

Cheshire J (1989). 'Great Britain.' *Sociolinguistica 3,* 42–53.

Cheshire J (1999). 'Spoken standard English.' In Bex T & Watts R J (eds.) *Standard English: the widening debate.* London/New York: Routledge.

Davies W V (2000a). 'Language awareness amongst teachers in a central German dialect area.' *Language Awareness 9,* 119–134.

Davies W V (2000b). 'Linguistic norms at school: a survey of secondary-school teachers and trainee teachers in a central German dialect area.' *Zeitschrift für Dialektologie und Linguistik 67,* 129–147.

Davies W V (2001). 'Standardisation and the school: norm awareness and norm tolerance in the educational domain.' *Linguistische Berichte 188,* 393–414.

DES (1989). *English for Ages 5–16* [the Report of the National Curriculum English Working Group, chaired by Professor Brian Cox]. London: Department of Education and Science.

Durrell M (1999). 'Standardsprache in England und Deutschland.' *Zeitschrift für germanistische Linguistik 27,* 285–308.

Fought C (2002). 'Californian students' perceptions of, you know, regions and dialects?' In Long D & Preston D (eds.) *Handbook of perceptual dialectology,* vol. 2. Amsterdam: John Benjamins.

Hartley L & Preston D (1999). 'The names of US English: Valley girl, cowboy, yankee, normal, nasal and ignorant.' In Bex T & Watts R J (eds.) *Standard English: the widening debate.* London/New York: Routledge. 207–238.

Heller M & Martin-Jones M (2001). *Voices of authority: education and linguistic difference.* Westport, CT/London: Ablex.

Jäger S (1980). 'Standardsprache.' In Althaus H-P, Henne H & Wiegand H E (eds.) *Lexikon der germanistischen Linguistik,* 2nd edn. Tübingen: Max Niemeyer. 375–379.

Jones M C (1998). *Language obsolescence and revitalization.* Oxford: Clarendon Press.

Lippi-Green R (1997). *English with an accent.* London: Routledge.

Milroy J & Milroy L (1999). *Authority in language* (3rd edn.). London/New York: Routledge.

Martin-Jones M & Saxena M (2001). 'Turn-taking and the positioning of bilingual participants in classroom discourse: insights from primary schools in England.' In Heller M & Martin-Jones M (eds.). 117–138.

Niedzielski N & Preston D (1999). *Folk linguistics.* Berlin: de Gruyter.

Pinker S (1994). *The language instinct. The new science of language and mind.* London: Penguin.

Townson M (1992). *Mother-tongue and fatherland: language and politics in German.* Manchester: Manchester University Press.
Trask R L (1999). *Key concepts in language and linguistics.* London: Routledge.
Wolff H (1959). 'Intelligibility and interethnic attitudes.' *Anthropological Linguistics 1,* 34–41.

Relevant Websites

www.phon.ucl.ac.uk/home/dick/lagb-las.htm – For matters relating to language and education in general.
www.lsadc.org – For matters relating to Ebonics.

Language Education: Grammar

R Hudson, University College London, London, UK

History

The earliest known evidence for grammar teaching dates to the early second millennium B.C. in Babylon, where scribes learned to write Sumerian and Akkadian, and to translate between the two, with the help of tablets that systematized the word-formation rules (Gragg, 1994). There is a great deal more evidence for grammar teaching from the end of the first millennium B.C. in classical Greece and in the world influenced by it, including Rome (Howatt, 1994). This is when the term *grammatiké* came into use, first for the understanding of letters (Greek *grammata*), and only later for what we call *grammar* (Robins, 1967: 13); the terminology reveals the intimate connection between the study of grammar and the teaching of writing. This was the start of the tradition of grammar teaching that persisted in Europe and the Near East through the next two millennia, and which was exported to much of the rest of the world.

In contrast, the grammatical tradition that started at about the same time (350 B.C.) in India was linked to religion rather than to the teaching of writing (Kiparsky, 1994). Here we focus on the European tradition of grammar teaching at school.

Grammar teaching was a central part of the school curriculum through the Middle Ages and beyond. At first, grammar supported the learning of Latin (as a second language), but later it was applied to the national languages such as English. In many countries, this tradition has continued uninterrupted to the present; this seems to be true, in general, of Eastern Europe and the Romance-speaking world. Like any other school subject, grammar needs renewal from the academic world, even if only through higher-level teaching of Latin or other ancient languages (including Old English); indeed, in at least some countries academic grammarians have seen schools as important users of their ideas. For example, the theoretical doyen of dependency grammar, Lucien Tesnière, arranged for his ideas to be trialed with schoolchildren (Tesnière, 1959: 4). In contrast, grammatical theorizing never had a serious place in the universities of England, and even in language learning it was in serious decline by the early 20th century, so there was no renewal, and teachers could only repeat what they themselves had learned at school. This is almost certainly one of the main reasons why grammar teaching disappeared from most schools in England from about 1930 to 1980, though it has now returned thanks to the National Literacy Strategy (Hudson and Walmsley, unpublished). School-level grammar teaching died at about the same time in other parts of the English-speaking world, and it is hard to predict future developments.

The Scope of Grammar Teaching

Grammar teaching at school always deals with the major word classes (the traditional parts of speech), but (at least in the English-speaking world) all too often it stops there. In principle, it can cover the full range of concerns that linguists call grammar; and indeed, in the absence of any other name for language structure, it can be taken even more broadly to include all aspects of language structure. In this very broad sense, therefore, school-level grammar teaching could include all the following topics:

- word classes, feeding all the other kinds of analysis
- syntax (including sentence and clause analysis)
- morphology (particularly important for English spelling)
- phonology, graphology, and spelling
- punctuation
- lexical relations (semantic and morphological relations)
- semantics (lexical and syntactic)
- pragmatics and discourse structure

Townson M (1992). *Mother-tongue and fatherland: language and politics in German.* Manchester: Manchester University Press.
Trask R L (1999). *Key concepts in language and linguistics.* London: Routledge.
Wolff H (1959). 'Intelligibility and interethnic attitudes.' *Anthropological Linguistics 1,* 34–41.

Relevant Websites

www.phon.ucl.ac.uk/home/dick/lagb-las.htm – For matters relating to language and education in general.
www.lsadc.org – For matters relating to Ebonics.

Language Education: Grammar

R Hudson, University College London, London, UK

History

The earliest known evidence for grammar teaching dates to the early second millennium B.C. in Babylon, where scribes learned to write Sumerian and Akkadian, and to translate between the two, with the help of tablets that systematized the word-formation rules (Gragg, 1994). There is a great deal more evidence for grammar teaching from the end of the first millennium B.C. in classical Greece and in the world influenced by it, including Rome (Howatt, 1994). This is when the term *grammatiké* came into use, first for the understanding of letters (Greek *grammata*), and only later for what we call *grammar* (Robins, 1967: 13); the terminology reveals the intimate connection between the study of grammar and the teaching of writing. This was the start of the tradition of grammar teaching that persisted in Europe and the Near East through the next two millennia, and which was exported to much of the rest of the world.

In contrast, the grammatical tradition that started at about the same time (350 B.C.) in India was linked to religion rather than to the teaching of writing (Kiparsky, 1994). Here we focus on the European tradition of grammar teaching at school.

Grammar teaching was a central part of the school curriculum through the Middle Ages and beyond. At first, grammar supported the learning of Latin (as a second language), but later it was applied to the national languages such as English. In many countries, this tradition has continued uninterrupted to the present; this seems to be true, in general, of Eastern Europe and the Romance-speaking world. Like any other school subject, grammar needs renewal from the academic world, even if only through higher-level teaching of Latin or other ancient languages (including Old English); indeed, in at least some countries academic grammarians have seen schools as important users of their ideas. For example, the theoretical doyen of dependency grammar, Lucien Tesnière, arranged for his ideas to be trialed with schoolchildren (Tesnière, 1959: 4). In contrast, grammatical theorizing never had a serious place in the universities of England, and even in language learning it was in serious decline by the early 20th century, so there was no renewal, and teachers could only repeat what they themselves had learned at school. This is almost certainly one of the main reasons why grammar teaching disappeared from most schools in England from about 1930 to 1980, though it has now returned thanks to the National Literacy Strategy (Hudson and Walmsley, unpublished). School-level grammar teaching died at about the same time in other parts of the English-speaking world, and it is hard to predict future developments.

The Scope of Grammar Teaching

Grammar teaching at school always deals with the major word classes (the traditional parts of speech), but (at least in the English-speaking world) all too often it stops there. In principle, it can cover the full range of concerns that linguists call grammar; and indeed, in the absence of any other name for language structure, it can be taken even more broadly to include all aspects of language structure. In this very broad sense, therefore, school-level grammar teaching could include all the following topics:

- word classes, feeding all the other kinds of analysis
- syntax (including sentence and clause analysis)
- morphology (particularly important for English spelling)
- phonology, graphology, and spelling
- punctuation
- lexical relations (semantic and morphological relations)
- semantics (lexical and syntactic)
- pragmatics and discourse structure

- sociolinguistics, style, and dialectology
- language change

However, syntax has always had a central place in language teaching, especially in the debate about whether grammar teaching improves writing.

Another variable in grammar teaching concerns the language to which it is applied: the pupils' mother tongue (L1) or a second or foreign language (L2). In principle, of course, it can be applied to either or both, and in an ideal world both languages will be treated in terms of the same framework of concepts and terminology. Cross-language teaching can work well or badly, according to how sophisticated the teacher's and pupils' grasp of grammar is. The recent history of the English-speaking world has seen much the same rejection of grammar teaching in L2 teaching as in L1 teaching, but the pendulum has now swung back in its favor as a result of a large body of supportive research (Norris and Ortega, 2000). However, this article will focus on L1 teaching.

A third variable in L1 teaching is whether the teaching is descriptive or prescriptive. It almost always involves teaching a standard language to learners who do not yet know it, but this can be presented descriptively as a complement to the local non-standard language – an additional language variety for use on public occasions. The reasons why linguists prefer descriptive teaching to prescriptive are not only social but also intellectual and pedagogical – prescriptive teaching is inherently dogmatic and stuffed with pseudological arguments which are a poor model for any growing intellect.

Grammar teaching could, in principle, be embedded in a much broader framework of work on language, including psycholinguistic questions about child language, speech processing, animal communication and sociological questions about how language perpetuates prejudices, power structures, and so on. This broad program would be a course in general linguistics in all but name, and it is being actively pursued under the name of Language Awareness (Hawkins, 1994).

The Aims and Methods of Grammar Teaching

Since its origins in classical Greece, the main aim of grammar teaching has been to support the learning of literacy skills – hence the etymological link with *gramma* 'written letter' and *graphein* 'to write'. However, it is important to remember that the grammar of any language is itself a complex system rather than a disconnected list of categories or facts, so even if the ultimate aim is to improve writing, the immediate aim has to be an understanding of the system of grammar itself. Merely teaching the names and definitions of the word classes by rote is unlikely to have any benefit for writing.

How is grammar supposed to improve writing? At its weakest, grammar teaching targets a prescriptive list of errors, but healthy grammar teaching encourages growth, even though this typically leads to temporary errors at growth points. Grammar teaching helps young writers to develop richer grammatical repertoires:

- by providing a metalanguage for discussion of new constructions and weaknesses in the children's writing.
- by providing a deeper understanding of how language works, including a more or less conscious awareness of the resources available for particular functions.
- by directing learners' attention to the form of what they read, in the hope that they will notice new patterns and remember them.

We shall consider below whether grammar can in fact produce these benefits.

However, the improvement of writing (and reading) is only one of the possible aims of grammar teaching. Here are some others:

- Providing a metalanguage for L2 learning, and perhaps even for explicit comparisons between L1 and L2 structures (Hawkins and Towell, 1996). This kind of cross-language comparison is one of the pillars of the Language Awareness movement mentioned earlier, but it also underlies the old idea that learning Latin was a good way to teach an English speaker to write better English.
- Teaching children, as part of a liberal education, about an important part of themselves. In this approach, grammar is taught, alongside other areas of language, on much the same basis as, say, history – something every citizen ought to know something about. This is another application of Language Awareness, mentioned above.
- Using grammar as a vehicle for teaching scientific method. A number of small-scale experiments have shown that children can learn to formulate and test hypotheses about grammar, and that even small amounts (e.g., two weeks) of this activity can produce measurable improvements in general scientific reasoning (Honda, 1994; Honda and O'Neil, 1993).

Similarly, a wide range of methods have been used for teaching grammar, ranging from didactic teaching by rote, with a question-and-answer dialogue that included 'parsing', to discovery-based learning in

which children induce grammatical generalizations from their own data. The methods used obviously depend heavily on the aims of the teaching, but even when the aim is to improve writing, there is a major choice about the timing of the grammar teaching in relation to writing activity. There are three options:

1. **Separate** teaching teaches grammar separately from writing so that topics in grammar are not related to topics in writing.
2. **Reactive** teaching teaches grammar as and when needed, an approach strongly favored under the title 'grammar in context' (Weaver, 1996). When taken literally, this approach generally means in fact that grammar is never taught because of the time needed to teach most grammatical concepts (Bullock, 1975).
3. **Proactive** teaching anticipates the needs of a class and teaches a particular point of grammar with immediate links to those needs. This has probably been the basis for all successful grammar teaching, such as the National Literacy Strategy in England (Anon, 2000). The advantage of this approach is to allow grammar teaching to have its own systematic structure (e.g., teaching about verbs before teaching about tenses) while still being closely integrated with writing.

These choices lead to further choices about methods – e.g., the choice of whether or not to teach and use sentence diagramming, and, if so, what kind of diagrams to use. These questions lead in one direction into pedagogy and in the another direction into linguistic theory.

Evaluating the Success of Grammar Teaching

Does grammar teaching work? In particular, does it improve writing and reading skills? Until recently, the received wisdom was that research had shown convincingly that it did not. The following summary is typical: "Formal grammar instruction appears to contribute nothing to the development of writing and reading skills" (Elley, 1994). It is certainly easy to find examples of grammar teaching that were unsuccessful (by any criteria), but this merely shows that grammar can be taught unsuccessfully. In contrast, there is also a great deal of solid research evidence that grammar teaching does in fact work if teachers are sufficiently well-informed about grammar and if the teaching is proactive (in the sense explained above in 3.), i.e., closely integrated with activity to which the grammatical instruction is relevant. The following examples are representative:

- The exercise known as "sentence combining" has been shown to improve writing (Mellon, 1969; O'Hare, 1973; Hillocks and Mavrognes, 1986). In this method, students combine a list of simple sentences into a single sentence, with or without explicit comment on the structures concerned. This activity combines focus on a specific aspect of grammatical form with an immediate writing activity. Its effectiveness is an important demonstration that grammar teaching does not, in fact, have to be 'in context,' in the sense of reacting to children's own writing.
- Instruction about specific aspects of spelling such as the possessive apostrophe and irregular past-tense forms produced immediate improvements in these areas, as measured by tests before and after the instruction (Bryant *et al.*, 2002, 2004; Nunes *et al.*, 1997a, 1997b). The areas of grammar chosen for these experiments are known to cause serious problems for English learners because the underlying rules are not easy to infer from examples, so it is important to know that explicit instruction can in fact help.
- Teaching about complex noun phrases helped 18-year olds to understand such phrases in reading (Chipere, 2003). The experiment was specifically designed to separate the effects of grammatical knowledge from those of working-memory differences, and showed that explicit instruction in grammar improved comprehension (as well as recall); in fact, it brought a 'low-academic-ability' group up to the same level of comprehension as a high-ability group who received no instruction.

However, it has to be admitted that the research basis for such work is disappointingly thin, and it is still not clear exactly what distinguishes successful and unsuccessful grammar teaching.

See also: Language Education: Correctness and Purism; Language Education: Language Awareness; Language Education: Teacher Preparation; Mother Tongue Education: Nonstandard Language; Mother Tongue Education: Standard Language; Second and Foreign Language Learning and Teaching.

Bibliography

Anon (1998). *The grammar papers*. London: Qualifications and Curriculum Authority.

Anon (1999). *Not whether but how. Teaching grammar in English at key stages 3 and 4*. London: Qualifications and Curriculum Authority.

Anon (2000). *Grammar for writing*. London: Department for Education and Employment.
Bryant P, Devine M, Ledward A & Nunes T (2002). 'Spelling with Apostrophes and Understanding Possession.' *British Journal of Educational Psychology 67,* 91–110.
Bryant P, Nunes T & Bindman M (2004). 'The Relations Between Children's Linguistic Awareness and Spelling: The Case of the Apostrophe.' *Reading and Writing 12,* 253–276.
Bullock S A (1975). *Language for life*. London: Department for Education and Science.
Chipere N (2003). *Understanding complex sentences: native speaker variation in syntactic competence*. London: Palgrave Macmillan.
Elley W (1994). 'Grammar Teaching and Language Skill.' In Asher R E (ed.) *The encyclopedia of language and linguistics*. Oxford: Pergmon Press. 1468–1471.
Ellis R (1998). 'Teaching and research: Options in grammar teaching.' *TESOLl Quarterly 32,* 39–60.
Grabe W & Kaplan R (1996). *Theory and practice of writing*. London: Longman.
Gragg G (1994). 'Babylonian grammatical texts.' In Asher R E (ed.) *The encyclopedia of language and linguistics*. Oxford: Pergmon Press. 296–298.
Hawkins E (1994). 'Language awareness.' In Asher R E (ed.) *The encyclopedia of language and linguistics*. Oxford: Pergmon Press. 1933–1938.
Hawkins R & Towell R (1996). 'Why teach grammar?' In Engel D & Myles F (ed.). *Teaching grammar: perspective in higher education*. London: AFLS and Centre for Information on Language Teaching. 195–211.
Hillocks G & Mavrognes N (1986). 'Sentence combining.' In Hillocks G (ed.) *Research on written composition: new directions for teaching*. Urbana, IL: NCTE. 142–146.
Honda M (1994). Linguistic inquiry in the science classroom. Ph.D. diss. Harvard.
Honda M & O'Neil W (1993). 'Triggering science-forming capacity through linguistic inquiry.' In Hale K & Keyser J (eds.) *The view from building 20: essays in linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press. 229–255.
Howatt A (1994). 'Language teaching: History.' In Asher R E (ed.) *The encyclopedia of language and linguistics*. Oxford: Pergmon Press. 2020–2027.
Hudson R (1992). *Teaching grammar. A guide for the national curriculum*. Oxford: Blackwell.
Hudson R (2001). 'Grammar teaching and writing skills: The research evidence.' *Syntax in the Schools 17,* 1–6.
Hudson R & Walmsley J (in press). 'The English Patient: English grammar and teaching in the twentieth century.' *Journal of Linguistics*.
Kiparsky P (1994). 'Paninian linguistics.' In Asher R E (ed.) *Journal of Linguistics*. 2918–2923.
Mellon J (1969). *Transformational sentence combining*. Urbana, IL: National Council of Teachers of English.
Noguchi R R (1991). *Grammar and the teaching of writing: limits and possibilities*. Urbana, Ill.: National Council of Teachers of English.
Norris J & Ortega L (2000). 'Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis.' *Language Learning 50,* 417–528.
Nunes T, Bryant P & Bindman M (1997a). 'Learning to Spell Regular and Irregular Verbs.' *Reading and Writing 9,* 427–449.
Nunes T, Bryant P & Bindman M (1997b). 'Morphological spelling strategies: Developmental stages and processes.' *Developmental Psychology 33,* 637–649.
O'Hare F (1973). *Sentence combining: improving student writing without formal grammar instruction. Research Report No. 15*. Urbana, IL: National Council of Teachers of English.
Robins R H (1967). *A short history of linguistics*. London: Longman.
Stubbs M (1986). *Educational linguistics*. Oxford: Blackwell.
Tesnière L (1959). *Éléments de syntaxe structurale*. Paris: Klincksieck.
Weaver C (1996). *Teaching grammar in context*. Portsmouth, NH: Heinemann.
Wyse D (2001). 'Grammar for writing? A critical review of empirical evidence.' *British Journal of Educational Studies 49,* 411–427.

Language Education: Language Awareness

P Garrett, Cardiff University, Cardiff, Wales, UK

The constitution of the Association for Language Awareness (ALA) defines LA as 'explicit knowledge about language and conscious perception and sensitivity in language learning, language teaching and language use.' Hence, LA has always extended into many areas, including translation (Bowker, 1999) and literature (Carter, 1999). Here we focus, first, on what is normally regarded as 'grassroots' and 'classroom' LA, centered on language learning and teaching, and, second, on pertinent broader social and applied contexts of language use, social attitudes, etc., and on some of the implications of globalization.

LA is generally seen as emerging in the 1970s, in the writings of Halliday (1971), and with the recom-

Anon (2000). *Grammar for writing*. London: Department for Education and Employment.
Bryant P, Devine M, Ledward A & Nunes T (2002). 'Spelling with Apostrophes and Understanding Possession.' *British Journal of Educational Psychology 67,* 91–110.
Bryant P, Nunes T & Bindman M (2004). 'The Relations Between Children's Linguistic Awareness and Spelling: The Case of the Apostrophe.' *Reading and Writing 12,* 253–276.
Bullock S A (1975). *Language for life*. London: Department for Education and Science.
Chipere N (2003). *Understanding complex sentences: native speaker variation in syntactic competence*. London: Palgrave Macmillan.
Elley W (1994). 'Grammar Teaching and Language Skill.' In Asher R E (ed.) *The encyclopedia of language and linguistics*. Oxford: Pergmon Press. 1468–1471.
Ellis R (1998). 'Teaching and research: Options in grammar teaching.' *TESOLl Quarterly 32,* 39–60.
Grabe W & Kaplan R (1996). *Theory and practice of writing*. London: Longman.
Gragg G (1994). 'Babylonian grammatical texts.' In Asher R E (ed.) *The encyclopedia of language and linguistics*. Oxford: Pergmon Press. 296–298.
Hawkins E (1994). 'Language awareness.' In Asher R E (ed.) *The encyclopedia of language and linguistics*. Oxford: Pergmon Press. 1933–1938.
Hawkins R & Towell R (1996). 'Why teach grammar?' In Engel D & Myles F (ed.). *Teaching grammar: perspective in higher education*. London: AFLS and Centre for Information on Language Teaching. 195–211.
Hillocks G & Mavrognes N (1986). 'Sentence combining.' In Hillocks G (ed.) *Research on written composition: new directions for teaching*. Urbana, IL: NCTE. 142–146.
Honda M (1994). Linguistic inquiry in the science classroom. Ph.D. diss. Harvard.
Honda M & O'Neil W (1993). 'Triggering science-forming capacity through linguistic inquiry.' In Hale K & Keyser J (eds.) *The view from building 20: essays in linguistics in honor of Sylvain Bromberger*. Cambridge, MA: MIT Press. 229–255.
Howatt A (1994). 'Language teaching: History.' In Asher R E (ed.) *The encyclopedia of language and linguistics*. Oxford: Pergmon Press. 2020–2027.
Hudson R (1992). *Teaching grammar. A guide for the national curriculum*. Oxford: Blackwell.
Hudson R (2001). 'Grammar teaching and writing skills: The research evidence.' *Syntax in the Schools 17,* 1–6.
Hudson R & Walmsley J (in press). 'The English Patient: English grammar and teaching in the twentieth century.' *Journal of Linguistics*.
Kiparsky P (1994). 'Paninian linguistics.' In Asher R E (ed.) *Journal of Linguistics*. 2918–2923.
Mellon J (1969). *Transformational sentence combining*. Urbana, IL: National Council of Teachers of English.
Noguchi R R (1991). *Grammar and the teaching of writing: limits and possibilities*. Urbana, Ill.: National Council of Teachers of English.
Norris J & Ortega L (2000). 'Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis.' *Language Learning 50,* 417–528.
Nunes T, Bryant P & Bindman M (1997a). 'Learning to Spell Regular and Irregular Verbs.' *Reading and Writing 9,* 427–449.
Nunes T, Bryant P & Bindman M (1997b). 'Morphological spelling strategies: Developmental stages and processes.' *Developmental Psychology 33,* 637–649.
O'Hare F (1973). *Sentence combining: improving student writing without formal grammar instruction. Research Report No. 15*. Urbana, IL: National Council of Teachers of English.
Robins R H (1967). *A short history of linguistics*. London: Longman.
Stubbs M (1986). *Educational linguistics*. Oxford: Blackwell.
Tesnière L (1959). *Éléments de syntaxe structurale*. Paris: Klincksieck.
Weaver C (1996). *Teaching grammar in context*. Portsmouth, NH: Heinemann.
Wyse D (2001). 'Grammar for writing? A critical review of empirical evidence.' *British Journal of Educational Studies 49,* 411–427.

Language Education: Language Awareness

P Garrett, Cardiff University, Cardiff, Wales, UK

The constitution of the Association for Language Awareness (ALA) defines LA as 'explicit knowledge about language and conscious perception and sensitivity in language learning, language teaching and language use.' Hence, LA has always extended into many areas, including translation (Bowker, 1999) and literature (Carter, 1999). Here we focus, first, on what is normally regarded as 'grassroots' and 'classroom' LA, centered on language learning and teaching, and, second, on pertinent broader social and applied contexts of language use, social attitudes, etc., and on some of the implications of globalization.

LA is generally seen as emerging in the 1970s, in the writings of Halliday (1971), and with the recom-

mendations of the Bullock Report (DES, 1975) in the United Kingdom, drawing attention to underachievement in L2s in British schools. Some more enterprising teachers were soon independently establishing local projects to increase pupils' knowledge of languages and language. Cameron (1993) refers to three types of LA program at this time: for pupils about to start L2 learning, for pupils in multilingual/multicultural classrooms, and as prelinguistics courses for pupils in higher classes. The National Council for Language in Education (NCLE) working parties (1978) and conferences (1981, 1985) brought valuable institutional support to these grassroots initiatives, and Hawkins (e.g., 1984) added an enduring theoretical framework, with proposals for LA as a 'bridge' between mother tongue L1 and L2 education, and between primary and secondary education, and also with broader social objectives, such as parenting. Further impetus came in 1989 with a British Association for Applied Linguistics Seminar on LA (an event organized to bring together language practitioners using the term for different purposes in a range of contexts), and in 1992 with the First International Conference on LA, at which the ALA and its journal *Language Awareness* were launched. ALA conferences have followed biennially at venues in the United Kingdom and overseas (to date, Canada, Ireland, Sweden, France, and Spain).

LA often has been identified with the United Kingdom, as above, but equivalent projects are increasingly found elsewhere. EU-funded projects involving networks across several countries have been particularly impressive – for example, the primary-school focused Evlang – 'l'éveil aux langues' project (Candelier, 1998). Labels used to translate LA have occasionally aroused debate, as they can carry subtle distinctions, even within a single context (e.g., Gnutzmann's [1997] list of German terms). Related English terms include *consciousness-raising* (CR) and *Knowledge about Language* (KAL). James (1996) proposes reserving CR for the identification of discrepancies between present knowledge and target knowledge, and employing LA for the metacognition of knowledge that one already possesses without previously realizing one had it. KAL has been employed in some U.K. government reports. Cameron (1993) points to the different origins of KAL and LA, and the confusion arising from assuming synonymy. These two terms in particular reflect the LA debate around the role of explicit learning and explicit knowledge in language learning.

LA is associated with discovery-focused pedagogy. Often, learners engage in small-scale investigations requiring reflection or talk about how languages work, how they are learnt, and how they themselves can best focus their own learning. Hence, the development of metalanguage (e.g., Jaworski *et al.*, 2004; Berry, 2005), awareness of strategies of learning and communication, critical evaluation of the process of language learning (e.g., Garrett and Shortall, 2002), and working toward more autonomy in learning and use (Little, 1997) also feature. Reflective approaches are extended to language teacher education (e.g., Wright and Bolitho, 1993; Walsh, 2003), and also teacher development. Hence, projects bridging 'the space between' English and Modern Foreign Language teachers have involved the teachers sharing reflections on their different perspectives and experiences (Pomphrey and Moger, 1999; Turner and Turvey, 2002).

Benefits attributed to LA are viewed across five broad and overlapping dimensions: performance, cognitive, affective, social, and power (James and Garrett, 1991). The performance dimension concerns whether knowledge gained from greater awareness of language facilitates improved language use and learning. The cognitive dimension relates LA to an 'awareness of pattern, contrast, system, units, categories, rules of language' (Donmall, 1985: 7), and to the development of an 'analytic competence' that extends beyond language learning: a role Hawkins (1984) suggests might earlier have been fulfilled by learning Latin. The affective dimension is usually considered in terms of attitudes, motivation and curiosity accruing from LA. Hawkins's proposal for LA as a bridging subject to address poor achievement is directed as much at this dimension as the performance one. The social dimension is generally seen in terms of social harmonization in contexts of language diversity, and of building better relations between ethnic groups. The power dimension aims at increased sensitivity and empowerment to counter the manipulative and oppressive use of language, and to understand and counter language ideology. LA work with its prime focus on this dimension is usually referred to as Critical LA (Fairclough, 1999). It will be clear that the success or failure of LA work, in terms of both short-term and more enduring impacts, needs to be judged on these dimensions. We now turn to consider LA in the wider social context, where the last three dimensions in particular feature strongly.

Although some LA projects in schools aim at changing attitudes toward languages and their speakers, other work has studied the broader social backdrop of attitudes and stereotypes itself. Whereas some of this survey work has drawn its data from teachers and students (e.g., Garrett *et al.*, 2003, in Wales; Lochtman *et al.*, 2004, in Brussels), other work has studied the views of nonlinguists generally, arguing (*inter alia*) '. . . the undeniable importance it has in the

language professional's interaction with the public' (Preston, 1996: 72). LA-relevant work also has focused on specific contemporary issues in which language and communication can play a crucial role, from designing teaching programs to counter physician-patient communication barriers in HIV/AIDS medical interviews (Singy and Guex, 1997), and (to extend LA to communication and discourse awareness) understanding the communication complexities of HIV/AIDS campaigns (e.g., Perloff, 2001), to attitudinal studies of ageism in communication, and of 'good communication' between generations (e.g., Williams *et al.*, 2004).

Cameron (2000) has linked this general contemporary concern with 'good communication' with processes of globalization, which is increasingly referred to in other recent LA work. Some concerns itself with the increasing mobility of large numbers of people and the impact on language choice and use, the values attached to them, and social identities, whether they are tourists (Jaworski *et al.*, 2003), or recently immigrated communities (Yelenevskaya and Fialkova, 2003, in Israel) or their children who have grown up in the new community (Jørgensen and Quist, 2001, in Denmark), or people returning to visit locations of earlier family generations (Wray *et al.*, 2003, in Wales). Other work has explored the implications of the claim that, with the weakening of traditional social distinctions of class, sex, and age under globalization, traditional language situations and notions of standard language varieties have come under attack. Kristiansen (2001), for example, finds attitudinal evidence of a splitting in two of Standard Danish, with a standard emerging for the media that differs from that valued in the educational system. But globalization also is viewed in terms of the commodification of discourse, the increasing marketization in our lives and institutions, including education, and of how our increasingly knowledge-based societies impact not only on our jobs but also on our personal relationships and identities across the lifespan. LA has a challenging role to play in such globalization processes, to better prepare us for pursuing our social identities, and for resisting organizational incursions into our everyday lives. To this end, Fairclough (1999) has argued the importance of promoting educational programs in critical discourse awareness.

See also: Bilingualism and Second Language Learning; Consciousness, Thought and Language; Discrimination and Language; Education in a Multilingual Society; Educational Linguistics; Folk Linguistics; Foreign Language Teaching Policy; Immigrant Language Education; Language Attitudes; Language Education: Correctness and Purism; Language Policy in Multilingual Educational Contexts; Language Teaching Traditions: Second Language; Migration and Language Planning; Migration and Language; Minorities and Language; Minority Languages: Education; Mother Tongue Education: Nonstandard Language; Mother Tongue Education: Standard Language; Motivation and Attitude in Second Language Learning; Multiculturalism and Language; Nonnative Speaker Teachers; Norms and Correctness; Politics, Ideology and Discourse; Second and Foreign Language Learning and Teaching.

Bibliography

Berry R (ed.) (2005). 'Metalanguage in applied linguistics.' Special issue of *Language Awareness 14,* 1–80.

Bowker L (1999). 'Exploring the potential of corpora for raising language awareness in student translators.' *Language Awareness 8,* 160–173.

Cameron D (2000). *Good to talk?: living and working in a communication culture.* London: Sage.

Cameron L (1993). 'Degrees of knowing: an exploration of progressive abstraction in language awareness work.' *Language Awareness 2,* 3–14.

Candelier M (1998). 'L'éveil aux langues á l'école primaire: le programme européen Evlang.' In Billiez J (ed.) *De la didactique des langues á la didactique du plurilinguisme: hommage á Louise Dabène.* Grenoble: CDL-Lidilem. 299–308.

Carter R (ed.) (1999). *Literature and language awareness.* Special issue of *Language Awareness 8,* 1–64.

DES (Department of Education, Science) (1975). *A language for life: the Bullock Report.* London: HMSO.

Donmall B G (ed.) (1985). *Language awareness.* London: CILT.

Fairclough N (1999). 'Global capitalism and critical awareness of language.' *Language Awareness 8,* 71–83.

Garrett P & Shortall T (2002). 'Learners' evaluations of teacher-fronted and student-centred classroom activities.' *Language Teaching Research 6,* 25–57.

Garrett P, Coupland N & Williams A (2003). *Investigating language attitudes: social meanings of dialect, ethnicity and performance.* Cardiff: University of Wales Press.

Gnutzmann C (1997). 'Language awareness: progress in language learning and language education, or reformulation of old ideas?' *Language Awareness 6,* 65–74.

Halliday M A K (1971). 'Foreword.' In Doughty P, Pearce J & Thornton G (eds.) *Language in use.* London: Edward Arnold. 4–6.

Hawkins E (1984). *Awareness of language: an introduction.* Cambridge: Cambridge University Press.

James C (1996). 'A cross linguistic approach to language awareness.' *Language Awareness 5,* 138–148.

James C & Garrett P (eds.) (1991). *Language awareness in the classroom.* London: Longman.

Jaworski A, Thurlow C, Lawson S & Ylänne-McEwen V (2003). 'The uses and representations of local languages

in tourist destinations: a view from British TV holiday programmes.' *Language Awareness 12,* 5–29.
Jaworski A, Coupland N & Galasiński D (eds.) (2004). *Metalanguage: Social and ideological perspectives.* The Hague: Mouton.
Jørgensen J N & Quist P (2001). 'Native speakers' judgements of second language Danish.' *Language Awareness 10,* 41–56.
Kristiansen T (2001). 'Two standards: one for the media, one for the school.' *Language Awareness 10,* 9–24.
Little D (1997). 'Language awareness and the autonomous language learner.' *Language Awareness 6,* 93–104.
Lochtman K, Lutjeharms M & Kermarrec G (2004). 'Duits en andere 'vreemde' talen in Brussel: een attitudenonderzoek bij Brusselse studenten uit een economische studierichting.' In Housen A, Pierrard M & Van de Craen P (eds.) *Taal, Attitude en Onderwijs in Brussel (Brusselse Thema's 12).* Brussel: VUB Press. 209–239.
Perloff R (2001). *Persuading people to have safer sex.* Mahwah, NJ: Erlbaum.
Pomphrey C & Moger R (1999). 'Cross-subject dialogue about language: attitudes and perceptions of PGCE students of English and modern languages.' *Language Awareness 8,* 223–236.
Preston D (1996). 'Whaddayaknow?: the modes of folklinguistic awareness.' *Language Awareness 5,* 40–74.
Singy P & Guex P (1997). 'Scope and limits of medical discourse concerning AIDS prevention.' *Language Awareness 6,* 238–241.
Turner K & Turvey A (2002). 'The space between shared understandings of the teaching of grammar in English and French to year 7 learners: student teachers working collaboratively.' *Language Awareness 11,* 100–113.
VanLier L (1995). *Introducing language awareness.* Harmondsworth: Penguin.
Walsh S (2003). 'Developing intercultural awareness in the second language classroom through teacher self-evaluation.' *Language Awareness 12,* 124–142.
Williams A, Garrett P & Tennant R (2004). 'Seeing the difference, feeling the difference: emergent adults' perception of communication and 'good' communication with peers and adolescents.' In Ng S H, Candlin C N & Chiu C Y (eds.) *Language matters: communication, identity, and culture.* Hong Kong: City University of Hong Kong Press. 111–136.
Wray A, Evans B, Coupland N & Bishop H (2003). 'Singing in Welsh, becoming Welsh: "turfing" a "grass roots" identity.' *Language Awareness 12,* 49–71.
Wright T & Bolitho R (1993). 'Language awareness: a missing link in teacher education?' *English Language Teaching Journal 47,* 292–304.
Yelenevskaya M & Fialkova L (2003). 'From muteness to eloquence: immigrants' narratives about languages.' *Language Awareness 12,* 30–48.

Language Education: Standards Movement

A Davies, University of Edinburgh, Edinburgh, UK

> Accountability has become a watchword in public sector enterprises. It is one of several themes . . . that are symptomatic of the late 20th century loss of confidence in the state as a provider of services. Nowhere have demands for accountability become more strident than in education (Moore, 2001: 177).

The current drive for accountability may explain the pervasiveness and ubiquity of standards and therefore delude us into the naïve view that the standards concept is new and original. It is not. The concept of and the concern for standards have a long tradition, sometimes under different names, the most common probably being norms, but there are other familiar terms, too, such as *rules* and *conventions*. What they all indicate is that there are social goals and that there are agreed-upon ways of reaching toward those goals.

In language studies, one of the most common uses of a standard is bracketed with a language, thus Standard Language (Standard English, Standard French, and so on). What is meant here is that one (or more) dialect(s) of a language is attached to social prestige and therefore accorded official status in education, publishing, teaching foreigners, and so on. Because of its official uses, descriptions of the standard written language are likely to be readily available in published grammars, books of usage, dictionaries, and style and punctuation manuals. What are unlikely to be so readily available are manuals on how to speak the standard language.

Standards, then, are ways of behaving, like conventions, but within institutions they have more authority because they can be used as nonnegotiable goals (Elder, 2000a). Brindley places standards under the broad heading of outcome statements: these, he considers, can refer to standards themselves and to benchmarks, attainment targets, bandscales, profiles, and competencies, all of which are "broadly speaking, standards of performance against which

in tourist destinations: a view from British TV holiday programmes.' *Language Awareness 12,* 5–29.
Jaworski A, Coupland N & Galasiński D (eds.) (2004). *Metalanguage: Social and ideological perspectives.* The Hague: Mouton.
Jørgensen J N & Quist P (2001). 'Native speakers' judgements of second language Danish.' *Language Awareness 10,* 41–56.
Kristiansen T (2001). 'Two standards: one for the media, one for the school.' *Language Awareness 10,* 9–24.
Little D (1997). 'Language awareness and the autonomous language learner.' *Language Awareness 6,* 93–104.
Lochtman K, Lutjeharms M & Kermarrec G (2004). 'Duits en andere 'vreemde' talen in Brussel: een attitudenonderzoek bij Brusselse studenten uit een economische studierichting.' In Housen A, Pierrard M & Van de Craen P (eds.) *Taal, Attitude en Onderwijs in Brussel (Brusselse Thema's 12).* Brussel: VUB Press. 209–239.
Perloff R (2001). *Persuading people to have safer sex.* Mahwah, NJ: Erlbaum.
Pomphrey C & Moger R (1999). 'Cross-subject dialogue about language: attitudes and perceptions of PGCE students of English and modern languages.' *Language Awareness 8,* 223–236.
Preston D (1996). 'Whaddayaknow?: the modes of folklinguistic awareness.' *Language Awareness 5,* 40–74.
Singy P & Guex P (1997). 'Scope and limits of medical discourse concerning AIDS prevention.' *Language Awareness 6,* 238–241.
Turner K & Turvey A (2002). 'The space between shared understandings of the teaching of grammar in English and French to year 7 learners: student teachers working collaboratively.' *Language Awareness 11,* 100–113.
VanLier L (1995). *Introducing language awareness.* Harmondsworth: Penguin.
Walsh S (2003). 'Developing intercultural awareness in the second language classroom through teacher self-evaluation.' *Language Awareness 12,* 124–142.
Williams A, Garrett P & Tennant R (2004). 'Seeing the difference, feeling the difference: emergent adults' perception of communication and 'good' communication with peers and adolescents.' In Ng S H, Candlin C N & Chiu C Y (eds.) *Language matters: communication, identity, and culture.* Hong Kong: City University of Hong Kong Press. 111–136.
Wray A, Evans B, Coupland N & Bishop H (2003). 'Singing in Welsh, becoming Welsh: "turfing" a "grass roots" identity.' *Language Awareness 12,* 49–71.
Wright T & Bolitho R (1993). 'Language awareness: a missing link in teacher education?' *English Language Teaching Journal 47,* 292–304.
Yelenevskaya M & Fialkova L (2003). 'From muteness to eloquence: immigrants' narratives about languages.' *Language Awareness 12,* 30–48.

Language Education: Standards Movement

A Davies, University of Edinburgh, Edinburgh, UK

> Accountability has become a watchword in public sector enterprises. It is one of several themes . . . that are symptomatic of the late 20th century loss of confidence in the state as a provider of services. Nowhere have demands for accountability become more strident than in education (Moore, 2001: 177).

The current drive for accountability may explain the pervasiveness and ubiquity of standards and therefore delude us into the naïve view that the standards concept is new and original. It is not. The concept of and the concern for standards have a long tradition, sometimes under different names, the most common probably being norms, but there are other familiar terms, too, such as *rules* and *conventions*. What they all indicate is that there are social goals and that there are agreed-upon ways of reaching toward those goals.

In language studies, one of the most common uses of a standard is bracketed with a language, thus Standard Language (Standard English, Standard French, and so on). What is meant here is that one (or more) dialect(s) of a language is attached to social prestige and therefore accorded official status in education, publishing, teaching foreigners, and so on. Because of its official uses, descriptions of the standard written language are likely to be readily available in published grammars, books of usage, dictionaries, and style and punctuation manuals. What are unlikely to be so readily available are manuals on how to speak the standard language.

Standards, then, are ways of behaving, like conventions, but within institutions they have more authority because they can be used as nonnegotiable goals (Elder, 2000a). Brindley places standards under the broad heading of outcome statements: these, he considers, can refer to standards themselves and to benchmarks, attainment targets, bandscales, profiles, and competencies, all of which are "broadly speaking, standards of performance against which

learners' progress and achievement can be compared" (Brindley, 1998: 48).

In language assessment, standards have three senses. I list them here and then discuss each in turn:

1. the skills and/or knowledge required to achieve mastery and the proficiency levels leading to mastery, along with the measures that operationalize these skills and/or knowledge and the grades indicative of mastery at each level.
2. the full set of procedures followed by test constructors which provide evidence to stakeholders that the test/assessment/examination/evaluation is serious and can be trusted, demonstrating, often through a code of ethics, that the test constructors are operating professionally.
3. a shorthand way of combining (1) and (2).

Standards as the Goal

Standards as the goal, the level of performance required or explained, thus "the standard required for entry to the university is an A in English"; "English standards are rising" (Davies *et al.,* 1999: 185). Stakeholders, of course, rightly wish to know what is meant by such statements, how they are arrived at, and what the evidence is for making them. For this understanding, there are three requirements: description, measurement, and reporting. The first is that the levels of performance (the 'standards') need to be made explicit. Such description is not easy (McNamara, 1997) and in second language learning becomes less and less easy as the learner progresses. In the initial stages, we might say: you must master these 50 vocabulary items. At later stages, we say, for example: you must demonstrate understanding of texts taken from given newspapers, facility to initiate and maintain a conversation with a native-speaking peer, the capacity to write a letter of complaint, and so on. These types of achievement are more difficult to delineate and to circumscribe. And even the initial level (50 vocabulary items) may, depending on our view of language learning, be absorbed into more complex uses of those items. But the three-stage process remains: there needs to be a description of the standard or level, an explicit provision of the measure that will indicate that the level has or has not been reached, and a means of reporting that decision, through grades, scores, impressions, profiles, and so on. Description, measure, and report; these three stages are essential, although there may be blurring of stages 2 and 3, such that the report is included within the measure. Where classical objective tests such as the Test of English as a Foreign Language (TOEFL) and the International English Language Testing System (IELTS) differ from the scale approaches of the Inter-Agency Round Table (IAR), the American Council on the Teaching of Foreign Languages (ACTFL), and the International Second Language Proficiency Ratings (ISLPR) is in their unequal implementation of the three stages. The tests offer measures and reports, but may be light on the first stage, description. The scales provide description and reports but may lack a measuring instrument. Proponents of the two methods (tests and scales) do not readily accept such criticism, maintaining that, in the case of the testing approach, the description is incorporated into the test and its supporting manuals; and in the case of the scaling approach, that the scale itself (which we consider to be a description) incorporates its own measuring instrument. *Quot homines, tot sententiae*. Ideally, scales require the support of a test instrument to provide reliable assignment to a level (or band) on a scale, in other words to offer an objective cutoff rather than leave the designation of a level to the subjective judgments of the interviewer/judge/rater Davies, 1995). However, it is not only scalar systems that are open to such claims of unreliability (Leung and Teasdale, 1997): tests such as IELTS incorporate components of scalar judgments in their Speaking and Writing subtests where performance is judged by only one rater, effectively offering a scalar outcome (and lending themselves open to questions about reliability).

The move away from the objective test to the subjective scale is no doubt part of the widespread rejection in the social sciences of positivism, fueled by the sociocultural turn and concern for critical language testing. But it also has a more practical explanation. In large scale operations, common standards may be more readily acceptable if they are imposed by a scale which is open to local interpretations (much the same may be said of an international code of practice). A contemporary example of this is found in the Council of Europe's Common European Framework (CEF) of Reference for Languages: Learning, Teaching, Assessment (CEF, 2003). The CEF claims to be the "basis for a coherent and transparent structure for effective teaching/learning and assessment relevant to the needs of learners as well as society, and that can facilitate personal mobility" (CEF, 2003: ix). Building on the earlier Threshold level and the Common Reference Levels (A1-C2), "the CEF is now . . . inspiring a new generation of sets of objectives for curriculum developers . . . . This current Manual, with its emphasis on relating assessments to one another through the mediation of the CEF, is a logical complement to the developments on levels and objectives' (CEF, 2003: ix). The CEF, then, is not a measure. For measuring purposes, the CEF operates as a common

reference to which local and national assessment instruments can relate (Taylor, 2004). This makes good sense, but there is undoubtedly the danger of the tail wagging the dog: that, to enable mediation through the CEF, local systems will massage their own measures so as to secure accord with the CEF. In other words, the apparent freedom offered by the CEF to local systems could be an illusion, and, as Fulcher (2004) insists, established without proper research.

Large-scale operations such as the CEF may be manipulated unthinkingly (rather than, as we have suggested, deviously) by juggernaut-like centralizing institutions. Mitchell describes the misconceived imposition of the Attainment Targets and Level Descriptors of the United Kingdom's National Curriculum for Modern Foreign Languages, asserting that the longer-term impact of these standards "will certainly be to reduce diversity and experimentation . . . we are likely to lose the more ambitious and more experiential interpretations of communicative language teaching, which has . . . historically been found at [the] local level' (Mitchell, 2001: 174). Elder reports a similar case of inappropriate standards for LOTE (Languages other than English in Australia (Elder, 2000b). Bailey and Butler, discussing the No Child Left Behind (NCLB) program in the United States, complain that, because of recent changes to the federal law, no distinction is made between English learners and native speakers. The law now requires "the inclusion of English Language Learner students in future mandated assessment systems. The NCLB Act of 2001 increases school accountability . . ." (Bailey and Butler, 2004). Such mismatches are not wholly unlike what we have suggested as the possible CEF massaging of local measures, because in all cases what is in train is the imposition of one overall set of standards nationally, regionally, or even universally, the McDonald'sization of language standards. However, our skepticism may be misjudged and out of place, because by their very nature standards are ambitious for wider and wider acceptance. There really is little point, after all, in establishing standards just for me if they have no meaning or application to you or anyone else: similarly with standards for a class, school, city, and so on. What then is wrong about the Mitchell, the Elder, and the Bailey and Butler cases is not that they were attempts at expanding the range and distribution of standards, but that they were the wrong standards for the populations discussed.

Setting Standards

If standards, in our first sense, refer to a level of performance required or experienced, the emphasis being on the language user or test-taker, a second sense refers to a set of principles which can be used as a basis for evaluating what language testers do, such as carrying out the appropriate procedures. When a school principal maintains that his/her school is 'maintaining standards,' the implication is that achievement levels over time are constant. When an examination body such as Educational Testing Service (ETS) or the University of Cambridge Local Examinations Syndicate (UCLES) claims that they are 'maintaining standards,' what they seem to mean is that they are carrying out the appropriate procedures, such as standard-setting (Griffin, 2001). Standard-setting is a technical exercise, involving, as it does, the determining of cut-scores for a test, either for pass/fail or for each level in a band system. But it is worth remembering, as Lumley *et al.* (1994) conclude, that standard setting remains a substantially political and ethical issue: "there can be no purely technical solution to the problem of standard setting in this context" (of an English test for ESL health professionals), the decision "remains intrinsically ethical and political; no amount of technical sophistication will remove the necessity for such decisions" (Lumley *et al.*, 1994: 39). Seeking to be ethical, behaving as responsible professionals, means that examination bodies must ensure that their products are reliable and valid, that they are properly maintained and renewed over time, that appropriate research is in place, and that the needs of all stakeholders are being addressed. It seems also necessary to demonstrate publicly that their claimed standards are being maintained. Hence, the professional statements embodied in a Code of Ethics (ILTA, 2000; ALTE, 2001). Skeptics may object that language testing was professional before testers felt the need to claim a professional status in a Code of Ethics, but there is another interpretation, that just as the reach of grammar expands into the area of discourse, so we are seeing in the professionalizing and ethicalizing of language testing a wider and wider understanding of validity.

Measuring and Reporting Standards

To an extent, this is where Messick's theorizing (1989) has taken us in his attempt to provide one overall coherent framework for the description, the measurement, and the reporting of standards, and the systematic effects they have on all stakeholders. The term that has come to be associated with his

conceptualization is that of *consequential validity*, but it does seem that *impact* may be an alternative name for it. Impact studies the effects that a test has when put to use: this is more than the more frequently used term *washback*, precisely because it is concerned not with just how a test works in one situation but with its systemic influences. As such, impact can investigate fundamental issues about standards: are they the right ones for the purposes intended, are they fully and openly described, are they attached to reliable and valid measures, and is the reporting clear and precise, and does the test produce desirable outcomes in the form of more appropriate and useful teaching. What impact studies, then, can do is to enable us to reevaluate and make explicit not just the standards we promote but the very view of language that we take for granted.

See also: Foreign Language Teaching Policy; Language Education: Correctness and Purism; Mother Tongue Education: Nonstandard Language; Native Speaker; Norms and Correctness; Second and Foreign Language Learning and Teaching.

Bibliography

ALTE (1994/2001). *The Association of Language Testers of Europe Code of Practice*. http://www.alte.org.

Bailey A L & Butler F A (2004). 'Ethical considerations in the assessment of the language and content knowledge of English language learners K-12.' *Language Assessment Quarterly 1(2, 3)*, 177–193.

Brindley G (1998). 'Outcomes-based assessment and reporting in language learning programmes: a review of the issues.' *Language Testing 15(1),* 45–85.

CEF (2003). *Relating language examinations to the common European Framework of Reference for Languages: learning, teaching, assessment (CEF). Manual: Pilot Preliminary Version.* Strasbourg: Council of Europe Language Policy Division.

Davies A (1995). 'Introduction: measures and reports.' *Melbourne Papers in Language Testing 4(2),* 1–11. University of Melbourne.

Davies A, Brown A, Elder C, Hill K, Lumley T & McNamara T (1999). *Dictionary of language testing.* Cambridge: University of Cambridge Local Examinations Syndicate and University of Cambridge.

Elder C (2000a). 'Preface.' In Elder C (ed.) *Defining standards and monitoring progress in languages other than English.* (Guest edited issue) *Australian Review of Applied Linguistics*, 23, 1–5.

Elder C (2000b). 'Learner diversity and its implications for outcomes-based assessment.' In Elder C (ed.) *Defining standards and monitoring progress in languages other than English.* (Guest edited issue) *Australian Review of Applied Linguistics, 23(2),* 36–61.

Elder C, Brown A, Grove K, Hill K, Iwashita N, Lumley T, McNamara T & O'Loughlin K (eds.) (2001). *Experimenting with uncertainty: essays in honour of Alan Davies.* Cambridge: Cambridge University Press.

Fulcher G (2004). 'Are Europe's tests being built on an unsafe framework?' *Guardian Weekly TEFL Supplement.* BBC World Service, March, 2004.

Griffin P (2001). 'Establishing meaningful language test scores for selection and placement.' In Elder *et al.* (eds.). 97–107.

ILTA (2000). *International Language Testing Association Code of Ethics.* http://www.dundee.ac.uk.

Leung C & Teasdale A (1997). 'Raters' understanding of rating scales as abstracted concept and as instruments for decision making.'. *Melbourne Papers in Language Testing 6(2),* 45–71. University of Melbourne.

Lumley T, Lynch B & McNamara T (1994). 'Are raters' judgements of language teacher effectiveness wholly language based?'. *Melbourne Papers in Language Testing 3(2),* 40–59. University of Melbourne.

McNamara T (1997). 'Wording the criterion statement: what's in a name?' *Plenary Address. National Working Forum on Assessment and Reporting in English Language and Literacy Programs: Trends and Issues.* Macquarie University: NCELTR.

Messick S (1989). 'Validity.' In Linn R L (ed.) *Educational measurement* (3rd ed.). New York: American Council on Education/Macmillan. 13–103.

Mitchell R (2001). 'Prescribed language standards and foreign language classroom practice: relationships and consequences.' In Elder *et al.* (eds.). 163–176. (2001).

Moore H (2001). 'Rendering ESL accountable: educational and bureaucratic technologies in the Australian context.' In Elder *et al.* (eds.). 177–190.

Taylor L (2004). 'Introduction: issues of test comparability.' *Research Notes 15:* UCLES. 2–5.

Language Education: Teacher Preparation

C Brumfit, University of Southampton,
Southampton, UK

Teacher education has always been closely bound up with the certification and recognition of qualified teachers, and language teacher education shares with other subject areas a close link to structures and practices laid down by governments. But two main characteristics distinguish this field from teacher preparation in other subject areas: first, it has been an unusually international practice as foreign language teachers tend to be more closely in touch with foreign ideas than their colleagues; second, the rise of English to its current position as the main language of international communication has led to a concentration of research and theory on untypical native-speaking teachers and the requirements of one particular market. Because of this demand, teacher education has taken place outside state-supported institutions as well as inside them.

This article surveys the history and current practices in second-language teacher education, using illustrations from a range of countries to characterize what is at the beginning of the 21st century a surprisingly integrated international practice. It uses as its baseline practices in Britain, because these have provided the most widely diffused models for the organization of teacher education.

The Origins of Teacher Education

Current models of teacher education reflect the export of Western European models of education to most parts of the world. Organized training of teachers is a product of the rise of mass education in Europe from the 19th century onward, though there were earlier roots in provision by religious groups for support of teachers in their schools, and some degree of apprenticeship of a less formal kind was common before the 19th century. At first, there was an assumption that advanced education could be left to the expertise of graduates with their advanced subject knowledge, but as state education extended, the need for cadres of primary school teachers (who were drawn from those with much lower professional qualifications) resulted in the establishment of systematic provision for the training of pupil teachers.

In the period following the Industrial Revolution, the pressure to improve the education of greater and greater numbers of young people was frustrated by the shortage of competent teachers. The Lancaster System, where the older pupils taught younger ones according to a carefully designed complex procedure laid down by Joseph Lancaster (1778–1838), was perhaps the most notable example of an attempt to pull education into some kind of administrative structure. Similar systems were in use in India by the end of the 18th century, and the systems of model schools and elementary training that were associated with these provided the foundations for the greater involvement of the state toward the middle of the 19th century.

In Britain, this was largely the product of a struggle between the established church and the government for control over the training of primary and infant teachers. Although Parliament voted money to support the founding of colleges, the vast majority of the 30 training colleges established by the middle of the 19th century were associated with the church. The curriculum embraced English grammar and literature, but no foreign or classical languages. With modifications, this provided a model for training which was accelerated after the 1870 Education Act introduced the idea of compulsory education. It was only in the same period that any training for secondary teachers began to be provided, again building on fragmentary work by religious bodies. Some of this was attached to universities, and a system of state-supported, but often religious-run, training for all levels persisted until the end of the 20th century.

The Emergence of Language Teacher Education

In the British tradition, foreign languages were not part of the primary curriculum, and until the 1960s were often considered suitable only for academically able students in the grammar schools and the independent fee-paying schools of the private sector. But such schools did not always expect their teachers to have training other than the subject-knowledge provided by their degree course. It was the introduction of comprehensive schooling, with all pupils from one area going to the same secondary school and receiving a variety of subject choices that encouraged the move to teacher training for all secondary teachers.

The rise of education as an organized higher education discipline is conventionally dated from the 1950s to the 1960s, though earlier provision, particularly in the United States, offered further sources of ideas and practice. The 1960s also saw a substantial expansion of general higher education provision. Up to that time, some secondary teachers had been

trained in universities, obtaining the one-year Postgraduate Certificate of Education (PGCE) following their degrees in their specialized subject, and then obtaining teaching posts in grammar schools. Separate colleges provided certification for primary teachers and for teachers in the non-academic secondary schools (Secondary Moderns).

Thus, teacher education was being redefined and recreated through the 1960s and 1970s. Colleges concerned with primary education tended to be staffed by holders of postgraduate Master's degrees from the universities in academic disciplines related to education, so the system became increasingly integrated on the new model. The main academic studies in these education programs were history, sociology, psychology, and philosophy of education. These were to accompany the practical competence arising from experience in schools ('teaching practice') and some understanding of the curriculum and teaching materials for trainees' own particular subject. But as the 1970s turned into the 1980s and the liberal consensus crumbled, governments stressed greater accountability and demanded greater control of the curriculum. So the academic subjects tended to be squeezed out in favor of work more closely directed to the content of the curriculum, which was increasingly subject to external government control. Skills in assessment (increasingly important when greater accountability was demanded) and a greater proportion of practical teaching experience was also insisted upon. At the same time, though, practices for secondary education varied from country to country. In mainland Europe, some countries, such as France, moved away from centralized models toward greater teacher autonomy at the same time as Britain was centralizing more substantially. Briefly, in the 1990s, there was government commitment to foreign languages for all learners, but that policy was retreated from as the new century arrived. In Britain, in 2004 all teachers were expected to 'deliver' a curriculum defined by the government, the details of which kept changing as a result of pressure for improvement. Such pressures result in government-sponsored initiatives that receive special funding and concentrate on particular areas of the curriculum such as literacy or numeracy. Language teacher education within Britain has been affected by the increased external control at the same time as governments have reduced their commitment to foreign language education. Overall, in state education, teachers in 2004 had reduced autonomy compared with 20 years before, and language teacher educators, whose funding is dependent on government sources, have to fit in with the policy of the government in power.

Language Teacher Education Since the 1960s

State Education: English-Speaking Countries

The big expansion in teacher education late in the 20th century coincided with a major expansion in linguistics as a discipline. In part, the claims that Chomskyan linguistics represented a decisive break with the past chimed well with the claims that teaching was entering a new and revolutionary state of expansion and influence. For language teachers, particularly, the 1960s seemed to herald a period of influence and usefulness. In Europe, the increasing success of the European Community (later European Union (EU)) gave a new emphasis to language learning, led by the influential publications of the Council of Europe which began to appear toward the end of the 1960s. Throughout the English-speaking world, initiatives through bodies such as the United States' Peace Corps and the British Council (for example, the 1962 Aid to Commonwealth English [ACE] program), heralded an expansion of English teaching throughout much of the world. The 1960s saw the foundation of international organizations for language teachers such as Teachers of English to Speakers of Other Languages (TESOL) (1966) and International Association of Teachers of English as a Foreign Language (IATEFL) (1967), of applied linguistics associations (BAAL, the British Association for Applied Linguistics, was founded in 1967), of bodies such as CILT (Centre for Information on Language Teaching and Research, 1966) in London, following the Center for Applied Linguistics (1959) in Washington. All were directly implicated in aspects of language teacher education.

Language teacher education in the state sector typically followed the pattern of teacher education for other subjects. Initial training followed one of two main patterns, being either concurrent with the learning of the main subject to be taught, or consecutive to this. Concurrent patterns, which enabled the main subject study to be closely geared to the needs of schools, would be found in Britain in B.Ed. degrees taking four years. Subject study was accompanied by educational study and by experience teaching in a school, largely based around a block placement over a period of several weeks or even months in the same school. In contrast, the consecutive pattern had the advantage of not separating teachers from other graduates in their main subject studies. Following completion of a normal degree program, a one-year PGCE program would offer educational theory, curriculum work, and teaching practice, sometimes followed by a probationary full-time period in a school before final

certification was awarded. This pattern offered the additional advantage that, in periods of difficult teacher supply, trainees did not have to choose education courses or an education career at the beginning of their university studies.

Initial teacher education was subject to control from local or central government, for it performed two complementary functions. Not only did it provide the training in the knowledge, skills, and attitudes necessary for effective teaching, but it also provided formal certification by the recognizing authority to practice as a teacher in state education. One effect of this dual role was that there remained constant tension between the autonomy of the higher education institution and the desire to influence the procedures and contents of the courses by those who funded and provided official recognition for the certificate. For this reason, structures were constantly changing, and the organization of initial teacher education rarely lasted more than a decade without some major restructuring being imposed. For language teaching, the situation was further confused by the power of the English as a Foreign Language (EFL) model. World English teaching was a powerful influence on teacher education in most languages throughout the second half of the 20th century. This was partly because of the power of the market, with publishers offering wide ranges of materials, some of which were truly innovative, and with the funding of major projects on grammars, dictionaries, and databases in response to the profit opportunities available in the EFL sector. But government initiatives and the international demand for higher-level training did create research opportunities within universities that were less easily available in educational work for most other subject areas.

The TESOL effect in English-speaking countries thus made language teacher education untypical. Private and public sector procedures influenced each other. While training to teach might be embedded in a state system, it might also be geared toward teaching outside the country of origin. The PGCE at the University of London Institute of Education, for example, emerged from the training of colonial education officers, and never had an entirely comfortable relationship with training for second-language teaching for minority language communities in the United Kingdom, and by the 1980s was training many people for private sector language work, or for careers in the British Council directing teaching activities.

The specific ESOL (English for Speakers of Other Languages) part of the course had broadly similar contents from course to course across continents. For example, a survey of the commonest elements in TESOL methods courses in 22 North American institutions (Long, 1983) found the following (number of courses in brackets):

1. Historical overview of teaching methods (19)
2. Teaching speaking skills (18)
3. Curriculum/syllabus design (18)
4. Teaching techniques (18)
5. Innovative/unconventional teaching methods (17)
6. Materials (17)
7. Teaching writing skills (16)
8. Teaching listening skills (15)
9. Teaching reading skills (15)
10. Teaching grammar (14)
11. Language testing (12)
12. Teaching vocabulary (11)
13. Teaching culture (10)
14. Linguistics and language teaching (10).

Similar lists could be found describing the methodology component of training courses in East Africa, India, Britain, and elsewhere any time between the late 1960s and the 1990s.

Throughout this period, attempts were made to bring language teacher education generally more closely into line with the increasing wealth of research findings on language acquisition and learning. One of the few surveys to attempt to cross the lines between training for different languages, including the teaching of first languages (Brumfit, 1988) argued that primary and secondary schools should develop practices that brought all languages together in a coordinated policy for language development. But in practice, the increasing pressure of government control meant that foreign languages were marginalized rather than strengthened and the motivation to coordinate all language activity disappeared.

State Education: Other Countries

Many of the academic pioneers of language education outside Europe contributed directly to teacher education. In Japan, for example, Palmer (1922) on general principles of language teaching and Hornby (1954) on sentence patterns published significant work used on training courses; in India, West (1953) developed major new directions, building on Thorndike's American work (1932) on vocabulary and word levels. In Sudan and Uganda, Bright (Bright and MacGregor, 1970) integrated much of the earlier work and responded to contemporary developments in linguistics. Throughout what became the British Commonwealth, substantial contributions were being made to the development of language teaching theory and practice with most of the books and materials aimed at the training of teachers. But foreign language teaching throughout the 20th century was an international activity (though as the work in Japan indicates,

such activity even early in the century was not confined to the future Commonwealth). Directly and indirectly all these figures were influenced by the European Reform Movement (Passy, 1899), while later international training reflected the influence of scholars such as Fries (1952) at Michigan. In the latter part of the century, international collaboration in the Council of Europe (Trim *et al.,* 1973) was a major influence on Languages for Specific Purposes and the Communicative Movement, while figures such as the Bulgarian Lozanov (1978) were significant in the discussion of Humanistic Approaches (Stevick, 1976).

The developments under British colonial influence largely reflected the structures of British teacher education, but until independence was looming, education systems tended to be for a relatively small elite (India and Pakistan providing the main exceptions), and often, especially for language teaching at the higher levels, to rely on expatriate native speakers of English. Colleges and universities in emergent Asia and Africa trained teachers, using mostly the concurrent model but occasionally, as in Uganda and Zambia, the PGCE structure, which in those countries trained both local and expatriate teachers. The Teachers for East Africa scheme, for example, funded jointly by Britain and the United States, trained local teachers together with British and Americans at Makerere in Kampala, and also recruited already trained teachers from overseas to supply the expanding education systems of Kenya, Uganda, and Tanzania in the 1960s, the period immediately before and after independence in those countries.

In European countries, although teacher education originated in the 18th and 19th centuries, it was not until late in the 20th century that training became compulsory for all teachers in most countries. Compulsion also increased self-awareness: the end of the century saw a massive increase of empirical research on processes both of teaching and teacher education. Some procedures such as 'action research' were explicitly aimed at a training model, seeing empirical study as primarily concerned with improving good practice within the teacher's own institution. Whatever research model predominated, training now became more closely tied to empirical findings, and there was an assumption that a researching approach might be desirable for good teachers. In Hungary, The Netherlands, and Lithuania, for example, a research project is a formal part of the portfolio for initial training.

At the beginning of the 21st century, European teacher education was predominantly carried out in universities, though in France training institutions were directly affiliated with the Ministry of Education. In most countries, such a ministry was directly responsible for recognizing teacher education qualifications, though perhaps through a quasi-independent body as in the United Kingdom and Hungary. The number of languages for which a teaching qualification could be obtained exceeded ten in Austria, Bulgaria, Czech Republic, Finland, and France, but English was the main training language offered. Courses were generally between three and five years in length if they were concurrent with the undergraduate program; postgraduate courses varied from one to two years in length. In many countries, intercultural communication was taught in addition to language skill-maintenance, but a minority of countries offered neither in the training program itself and relied on the degree main subject courses for such provision. Some programs in, for example, Austria, France, and Germany provided opportunities for study abroad as part of the course. While all courses addressed language teaching methodology, with almost all of them including information technology, only about two-thirds provided explicit metalinguistic study for teachers. Assessment procedures varied widely, and might be, at least partially, by external state examination (e.g., Austria, France, Germany), by public lecture (French-speaking Belgium), by competitive examination (Romania), by separate language proficiency examinations (Austria, Hungary), by attitude and participation in the courses (Ireland, Norway, Sweden). Almost all countries included a block teaching practice (Greece and Cyprus providing exceptions), while some had official training schools attached to universities (Austria, Finland). Course content centered on communicative teaching methodology, but varying emphases on such issues as the role of the Internet, distance learning, Content and Language Integrated Learning (CLIL), and bilingual teaching ensured that distinctive national and institutional differences remained.

To give an example of a specific program, the Spanish *Certificado de Aptitud Pedagogica* (CAP) ran from October to April in the Education Institute of each university. In 1996 from October to January general theory was provided either by correspondence or by about 120 contact hours' attendance. In February and March, about 60 hours of practice in state secondary schools was provided. A short dissertation was submitted by the end of May which provided a description of the school, a detailed account of the subject grouping within which the student had taught, and a detailed account of two lessons taught.

In most countries, then, pre-service language teacher education is constrained by the structure of general teacher education but shares a strong internationally determined agenda, so that the assumptions of

communicative language teaching, for example, are, at least in principle, applied to many different languages. It is probably fair to say that language teachers are trained as teachers (who teach language), rather than as language specialists (who are teachers). At post-experience, in-service level greater concentration is possible, but the proportion of state teachers who regularly attend long advanced courses has never been high, and the best-organized in-service programs, as in Germany, have tended to concentrate on short courses.

Private Sector Provision

Private sector provision of teacher education presents a much more varied picture, for it is subject only to the discipline of the owners of private language schools. It is a fair generalization that most private language schools work with either untrained or state-trained teachers, but the international English teaching profession, particularly in the last two decades of the 20th century has worked hard to offer a clear route for qualification even among those teachers who do not have state qualifications. Private sector qualifications were generally distinguished by their concentration on language teaching skills to the exclusion of the more general education theory and background provided by state training.

A pioneer in this field was the British-based group, International House, led by its Director, John Haycraft (1926–1996). He initiated in the 1960s and 1970s an intensive short induction course for would-be teachers which integrated the processes of teaching and guidance from experienced teachers, working together in the language school, usually with volunteer students who received free extra classes. With the help of many creative teachers, he devised a distinctive style of language teaching, relying on exciting and active oral work with smaller classes than would be found in state institutions. Even state-trained teachers needed the adjustment provided by these courses as they shifted their focus to adults in smaller classes. Private language schools, many of whom ran similar courses as English language teaching expanded, fed their trainees into a national range of qualifications set up initially by the Royal Society of Arts (RSA) and later incorporated into the portfolio of the University of Cambridge Local Examinations Syndicate (UCLES), which had for years run language examinations for Commonwealth schools and for the international EFL sector. The RSA ran a part-time diploma program, the equivalent of about ten weeks full-time, which corresponded approximately to the subject-specialist part of a state PGCE. Another London-based body, Trinity College, runs a similar course, and a range of more-or-less specialized offshoots of these diplomas, for teachers of young children, for multicultural education, for non- native-speaking teachers, etc., have developed to cater to the expanding demand. In practice, the ten-week equivalent certificate has come to be recognized as the basic initial teaching qualification, though teachers who wish to progress in their careers have often moved on to university diplomas and masters courses to extend and deepen the more theoretical components of their initial training.

To illustrate the range of provision that was developing, a snapshot of the early 1980s may be helpful. At that time, courses of the initial four-week type, aimed at English native-speakers, were found in 19 centers in Britain and 3 overseas. In addition, many universities and colleges ran short summer courses of similar length.

The RSA 10-week diploma was offered at 51 centers in Britain and 41 overseas. In addition 26 centers in Britain offered the RSA Diploma for teaching English as a Second Language in the United Kingdom, aimed at settled minority language speakers. An unpredictable number of colleges of education and universities also offered a limited option in EFL as part of their conventional state education training (though these did not survive the tightening of central control over the teacher education that emerged over the following decade). An official state-supported PGCE in EFL/ESOL could be taken at universities in Aberystwyth, Bangor, Leeds, Leicester, London (both at the Institute of Education and St. Mary's College, Twickenham), Manchester, and Sheffield. Courses of this kind were all closed when the government withdrew funding in the early 1990s, but many of their graduates had careers at home and abroad weaving in and out of public and private sector institutions.

In the late 1980s, there was substantial pressure to rationalize this diversity and proliferation of courses. The term *Certificate in EFL* could embrace anything from a two-day course from an unrecognized language school to a one-year full-time course at a university. While language schools had established bodies for recognition of basic standards, this had not prevented many schools from thriving outside ARELS (Association of Recognised English Language Schools), and many attempts to establish control of teacher education qualifications failed to achieve consensus. BATQI (British Association of Teacher-Qualifying Institutions) started in 1990, but neither it nor its partial successor the British Institute of English Language Teaching successfully brought together public and private sectors into an agreed-upon framework. It has been left in effect for the Cambridge Examinations to provide the *de facto*

backbone for initial training in the private sector, with courses including Certificates and Diplomas for Teaching EFL to Adults and for Overseas Teachers of English, and the Cambridge Examination in English for Language Teachers. Initial training for ESOL is no longer available outside the private sector.

Continuing Professional Development

This article has so far concentrated on the initial training of language teachers, though reference has been made to post-experience provision when relevant. Whereas initial training has a fairly definable curriculum, with much consensus on what is required of competent teachers (in terms of classroom management, curriculum, and assessment needs, etc.), and specifically of language teachers (in methodology and materials), the position with continuing professional development is much more varied. Such provision performs a range of different functions:

1. to augment or substitute for unsatisfactory or non-existent initial teacher education
2. to update teachers' professional knowledge by introducing developments since their initial training
3. to extend the repertoire of pedagogic skills to fields for which they were not initially trained (e.g., heading a department, organizing an IT provision, etc.)
4. to develop individuals' personal education, enabling them to make more coherent sense of their accumulated professional craft-knowledge
5. to allow opportunities for the development of specific curricula or teaching materials that are needed for a changing professional scene
6. to reward loyal and long-serving teachers with time to reflect and develop.

Some of these provisions may be combined, but it is clear that the different needs of employers and the state, of individuals for personal development, and of the teaching institution from which individuals come may sometimes conflict. Arguments for in-service provision have often centered on which of these needs justifies the expenditure of scarce funding.

In many countries, continuing professional development saw an expansion in the 1970s, as educational fashions passed rapidly from country to country, particularly within supranational bodies such as the EU. In-service provision sometimes consolidated the work of teachers' associations that had been active for years before, as in Spain. Elsewhere, provision was laid down in hope (Portugal passed a law in 1973 stating that teachers had a right to continuing professional development provided by the state, but the revolution a year later meant that such provision did not begin to appear until after 1986).

Nonetheless, whether it is to an in-school, half-day class or a local government-organized, one-day conference, a week with a government-funded advisory group, or a full year or more to complete a taught Master's program, some sort of post-experience provision is likely to be within the reach of many language teachers. With regard to short courses (up to one month long), it is worth noting that there is a thriving private-sector-based market in introductions to new and exciting methods of teaching, sometimes drawing upon recognized traditions, such as humanistic language teaching, but more often relying on adaptation of ideas derived from these and other sources by institutions (for example the Pilgrims School in Canterbury) who have made a reputation as exciting innovators.

In some countries and regions, continuing professional development has been built in to the official career progression route. A survey of provision in a number of countries published in 1994, for example, reported that in most states of the United States teachers had to accumulate 150 credits every five years in order to have their license to teach renewed. Credits could be obtained from evening courses, summer seminars, lectures, workshops offered by professional associations, or by universities.

Longer courses, and particular Master of Arts programs, require slightly more extended comment, not least because they provide an internationally recognized qualification which is widely used as a prerequisite for entry to management, teacher education, inspectorate, and higher education posts. Such courses perform different functions in different societies. Countries that provide postgraduate initial teacher education sometimes attach such training to Master's degrees, and there are currently moves in several universities in Britain for parts of the PGCE to be given M.A. degree equivalence.

Historically, though, M.A. courses in Applied Linguistics or English Language Teaching (ELT) in the English-speaking world have concentrated on the following broad areas:

1. general theory, history of linguistics, etc.
2. general issues of language description (e.g., syntax, semantics, discourse analysis, pragmatics, phonology/phonetics, writing)
3. pedagogical issues (e.g., communicative language teaching, curriculum design, assessment, languages for specific purposes, information technology and language learning, teaching materials)

4. sociolinguistics/sociology of language (e.g., sociolinguistics, language and power, English as a world language, language policy, language and gender)
5. psychology and language (e.g., psycholinguistics, first and second language acquisition and learning).

Courses addressing initial teacher education concerns may include a practicum, but many courses expect several years of prior teaching experience before they admit students to the course.

Relationship with Research, etc.

Attitudes toward research have shifted radically as teacher education has become more established. Even for academics, research in the mid-20th century often related to the foundation disciplines rather than directly to the practice of teaching. Thus philosophers, historians, and psychologists of education practiced research, but those directly concerned with teaching in classrooms were far less likely to do so. The research that directly impinged on language teaching might derive from descriptive work on grammar and vocabulary studies but it rarely addressed pedagogic procedures in detail. Partly this was a result of the limitations of available technology. Observation schedules such as Flanders Interaction Analysis were cumbersome to use and subjective until the advent of easy audio- and video-recording. But there was also developing increased dissatisfaction with the impressionistic nature of much methodological discussion. Influential textbooks such as Billows' *The Techniques of Language Teaching* (1961), as well as the practice in many institutions, reported on the ideas and hunches of successful and often inspiring teachers. Attempts to systematize the whole process appeared (Mackey, 1965), but extensive empirical studies of practice in the normal classroom had to wait until the 1970s to become widespread. Only in the last quarter of the century were practitioners able to begin the process of basing their interpretations of successful practice on evidence drawn from observations in a variety of different settings. It is fair to observe that, as science, language teaching pedagogy is still in its infancy, for there simply have not been enough studies in a wide enough range of settings for reliable comparisons to be made between practices with different types of learner, teacher, or educational setting. While it is true that both teaching and teacher education now have stronger empirical bases than in the past, there has been a failure to replicate, or even associate, local studies with those elsewhere, which has limited the effectiveness of the studies that have been produced. It will be the work on teacher thinking, on classroom practices, and on processes of teacher learning and development that will determine the structure of teacher education in the next few decades.

See also: Applied Linguistics: Overview and History; Assessment of Second Language Proficiency; Communicative Language Teaching; Culture in Language Teaching; Foreign Language Teaching Policy; Fries, Charles Carpenter (1897–1967); Hornby, Albert Sidney (1898–1978); Language Education: Standards Movement; Language Teaching Traditions: Second Language; Nonnative Speaker Teachers; Passy, Paul Edouard (1859–1940); Second and Foreign Language Learning and Teaching.

Bibliography

Billows L (1961). *The Techniques of Language Teaching.* London: Longman.

Bright J & McGregor G (1970). *Teaching English as a Second Language.* London: Longman.

Brumfit C J (ed.) (1988). *Language in Teacher Education.* Brighton: National Congress on Languages in Education.

Byram M (1994). 'Training language teachers: views from elsewhere.' *Language Learning Journal 10,* 6–8.

Dent H C (1977). *The Training of Teachers in England and Wales 1800–1975.* London: Hodder & Stoughton.

Faber P (ed.) (1999). *English teacher education in Europe: new trends and developments.* Frankfurt: Peter Lang.

Fanselow J F & Light R L (eds.) (1977). *Bilingual, ESOL and foreign language teacher preparation: models, practices, issues.* Washington, DC: TESOL.

Finocchiaro M (1965). *English as a second language: from theory to practice.* New York: Regents.

Fries C C (1952). *The structure of English.* New York: Harcourt Brace.

Grenfell M, Kelly M & Jones M (2003). *The European language teacher: recent trends and future developments in teacher education.* Bern: Peter Lang.

Hornby A S (1954). *Guide to patterns and usage in English.* London: Oxford University Press.

Howatt A P R (2004). *A history of English language teaching* (2nd ed.). Oxford: Oxford University Press.

Long M (1983). 'Training the language teacher as researcher.' In Alatis J E, Stern H H & Strevens P (eds.) (1983). *Applied linguistics and the training of language teachers.* Washington, DC: Georgetown University Press. 281–297.

Lozanov G (1978). *Suggestology and outlines of suggestopedy* (original in Bulgarian, Sofia, 1971). London: Gordon and Breach.

Lugton R C (ed.) (1970). *Preparing the EFL teacher: a projection for the '70s.* Philadelphia: Center for Curriculum Development.

Palmer H E (1922). *The principles of language study.* London: Harrap.

Passy P (1899). *De la methode directe dans l'enseignement des langues vivante*. Paris: Colin.
Richards J C & Nunan D (eds.) (1990). *Second language teacher education*. Cambridge: Cambridge University Press.
Rivers W (1968). *Teaching foreign language skills*. Chicago: Chicago University Press.
Stevick E (1976). *Memory, meaning and method*. Rowley Mass: Newbury House.
Thorndike E L (1932). *A teacher's word book of 20,000 words*. New York: Teachers' College, Columbia University.
Trim J L M, Richterich R, van Ek J A & Wilkins D A (1973). *Systems development in adult language learning*. Strasbourg: The Council of Europe. (Reprinted 1980, Oxford: Pergamon Press.)
West M (1953). *A general service list of English words*. London: Longman.

Language Education – Vocabulary

P Nation, Victoria University of Wellington, New Zealand

What Vocabulary?

There are numerous vocabulary counting programs that change texts into frequency lists of vocabulary and show how much text is covered by frequency-ranked lists. What these counts show is that a relatively small number of words account for a very large proportion of text. The 270 function word types (176 word families), such as *one, a, the, because, in, must*, cover about 44% of the running words (tokens) found in most texts. These function words, however, make up a small proportion of the frequent words of English. Among the highly ranked items are words such as *time*, *say*, *day*. The most frequent 2000 words of English, which include most of the function words, are the essential widely used words of the language. They are important no matter what use is made of the language, and they cover a large proportion of spoken and written text. They also represent a feasible learning goal for an English course of 800 to 1000 hours.

For learners with academic purposes such as study through the medium of English in senior high school or university, the next important vocabulary learning goal is academic vocabulary. The best researched list of these words is the Academic Word List (Coxhead, 2000), which consists of 570 word families arranged into ten sublists with sublist 1 containing the most frequent widest-range items. There are well-researched tests available (Schmitt *et al.,* 2001; Nation, 2001) called the Vocabulary Levels Tests, which can be used to quickly determine which level of vocabulary learners need to focus on. The main distinction to be made when interpreting the results of the test is between the high frequency words (the first 2000 words and the Academic Word List) and the low frequency words (3000 level onward).

How Should Words Be Dealt With?

The high frequency/low frequency distinction is an important one when planning a vocabulary program. This is because from a teaching perspective these two groups of words should be dealt with in different ways. The 2000 or 2570 high frequency words deserve classroom time. This is because of their high frequency and wide occurrence. This classroom time can involve direct teaching of these words, doing activities designed to teach and practice these words, and deliberate learning of these words. Low frequency words on the other hand do not deserve classroom time. This is because there are so many of them and because of their low frequency of occurrence. When learners know the high frequency words, they need to begin learning low frequency words. However, the teacher should not be spending valuable classroom time teaching those low frequency words. Instead, the teacher should spend time teaching the four most useful strategies for dealing with these words, namely the strategies of guessing words from context clues, deliberate learning using word cards, helping memory by using word part analysis, and dictionary use.

The opportunities for teaching and learning high frequency vocabulary can be divided into four strands. The meaning-focused input strand involves opportunities for learning through listening and reading where learners are already familiar with 95–98% of the running words in the listening and reading input. The meaning-focused output strand involves opportunities for learning through speaking and writing. The language-focused learning strand involves giving deliberate attention to vocabulary learning through activities such as intensive reading, using word cards, preteaching of vocabulary before doing communicative activities, and the deliberate learning of vocabulary strategies. The fourth strand is the fluency development strand where learners do not learn new vocabulary but practice making the best use of words

Moir J & Nation I S P (2002). 'Learners' use of strategies for effective vocabulary learning.' *Prospect 17,* 15–35.
Nagy W E, Herman P & Anderson R C (1985). 'Learning words from context.' *Reading Research Quarterly 20,* 233–253.
Nation I S P (2000). Learning vocabulary in lexical sets: dangers and guidelines. *TESOL Journal 9,* 6–10.
Nation I S P (2001). *Learning vocabulary in another language.* Cambridge: Cambridge University Press.
Nation P & Wang K (1999). 'Graded readers and vocabulary.' *Reading in a Foreign Language 12,* 355–380.
Palmer D M (1982). 'Information transfer for listening and reading.' *English Teaching Forum 20,* 29–33.
Read J (1998). 'Validating a test to measure depth of vocabulary knowledge.' In Kunnan A J (ed.) *Validation in language assessment.* Mahwah, NJ: Lawrence Erlbaum. 41–60.
Read J (2000). *Assessing vocabulary.* Cambridge: Cambridge University Press.
Schmitt N & McCarthy M (eds.) (1997). *Vocabulary: description, acquisition and pedagogy.* Cambridge: Cambridge University Press.
Schmitt N, Schmitt D & Clapham C (2001). 'Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test.' *Language Testing 18,* 55–88.
Smith H (1996). 'An individualised vocabulary programme.' *TESOLANZ Journal 4,* 41–51.
Waring R & Takaki M (2003). 'At what rate do learners learn and retain new vocabulary from reading a graded reader?' *Reading in a Foreign Language 15,* 130–163.
Wesche M & Paribakht T S (1996). 'Assessing second language vocabulary knowledge: depth versus breadth.' *Canadian Modern Language Review 53,* 13–40.

Language Families and Linguistic Diversity

M Ross, The Australian National University, Canberra, Australia

Linguistic Diversity

'Linguistic diversity' refers to several interrelated phenomena:

1. phylogenetic (genetic, genealogical) diversity: the number of language families in a geographic area;
2. intrafamilial and intralinguistic variation: differences among the languages of a family and among the variants of a language;
3. language density: the number of languages in a geographic area;
4. typological diversity: differences among the structures which make up languages.

The focus of this article is on (1), (2) and (3) (*see also* **Variation and Language: Overview**). More conventional approaches to these topics are discussed first, then a recent alternative.

Although (4) will not be discussed further here, the relationship between phylogenetic and typological diversity should be mentioned. It is quite common for the members of a language family to be typologically similar, but there is no necessary reason why they should be so. There is, for example, considerable typological variety among the grammars of the Austronesian family, some of it the result of contact with non-Austronesian languages, but much of it the outcome of language-internal changes. Conversely, languages belonging to different language families may be typologically quite similar, either as the result of contact or as the result of independent parallel innovation.

Language Families

Definitions and Problems

A language family is conventionally defined as a set of languages that share a common ancestor. The family metaphor captures two insights: that languages are systems that are transmitted from one generation to the next (generational continuity) and that these systems change over time so that when a community of speakers divides into, say, three separate communities, the speech of each of the three will change in different ways from the others, leading eventually to mutual unintelligibility, i.e., three languages descended from a shared ancestor. The basis of the metaphor is biological evolution, where one species divides into two or more new species (McMahon, 1994: Chap. 12), even though linguists often use the term 'family tree,' which otherwise denotes trees drawn to represent human family relationships.

Identifying a language family in the first place depends on finding individual-identifying evidence (Nichols, 1996, 1997), patterned similarities between a set of languages that could not have arisen by chance and must be the outcome of shared inheritance. The Indo-European family, for example, was first definitively identified by Sir William Jones in 1786 on the basis of similarities between the verb paradigms of Sanskrit, Greek, and Latin. Recently, the Trans-New

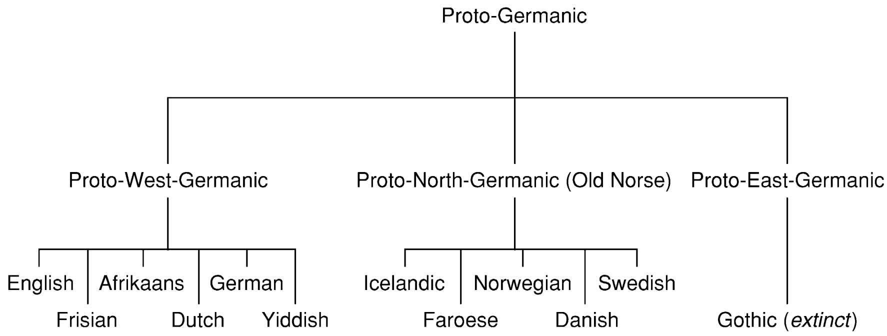


Figure 1 A conventional family tree of the Germanic languages.

Guinea family has been identified on the basis of formally similar pronouns (Ross, 2005).

Figure 1 shows a conventional family tree of the Germanic languages (*see* **Germanic Languages**) and serves as a starting point for a discussion of the family-tree model. It shows that the family concept is recursive. Thus English, Friesian, etc. form a West Germanic family. This, together with the North Germanic family and the single-member East Germanic family, forms the Germanic family. The latter, along with the Romance, Balto-Slavic, Indo-Iranian, and other, some single-member, families, is part of the Indo-European family (not shown in Figure 1).

The languages at all nodes other than the terminals in Figure 1 are reconstructed proto-languages. The changes that languages undergo are often regular enough to allow quite a safe reconstruction of their parent based on the correspondences among them, provided that the linguistic comparative method is carefully applied. Thus the reconstruction of Proto-Indo-European is reasonably secure. The reconstruction (and indeed the existence) of its hypothesized parent, Proto-Nostratic, however, is considered dubious by many linguists, as it is not based on a rigorous application of the comparative method (Nichols, 1997), the maximum reach of which is usually reckoned to be around 6000–8000 years (*see* **Long-Range Comparison: Methodological Disputes**).

The comparative method provides a means for inferring the tree structure of a large family by reconstructing the chronology of innovations relative to the protolanguage. For example, all Germanic languages reflect the so-called Germanic sound shift, a set of changes in consonants relative to the reconstructed system of Proto-Indo-European. It is inferred that this set of changes only occurred once, in Proto-Germanic, and this allows us to insert the Proto-Germanic node into the Indo-European tree (*see* **Subgrouping Methodology**).

Scholars investigating worldwide linguistic diversity need standardized phylogenetic units to work with. After Morris Swadesh devised the techniques known as lexicostatistics and glottochronology in the 1950s (Swadesh, 1972), there was a wave of optimism about the possibility of quantifying linguistic diversity. Researchers drew trees based on the percentage of putative cognates (words related through shared inheritance, e.g., English *house* and German *Haus*) in each pair of languages on Swadesh's 100- or 200-meaning list, setting percentage ranges for different levels of grouping.

For example, in work on Papuan languages two languages were deemed to be dialects of the same language if the difference between them meant no less than 70% of list items were cognate; at the other extreme lists with 5–12% cognates were attributed to the same phylum. The basic classification was into dialects, languages, subfamilies, families, stocks, and phyla. Today many scholars think such findings have little utility for historical reconstruction, as cognate pairs were often identified on the basis of similarity in form, with no attention to whether the words displayed regular correspondences determined by the comparative method, and cognacy was often not distinguished from borrowing. Wordlists could easily contain gaps and elicitation errors, skewing results. This meant that only lower-level groupings – those in any case obvious by inspection – were reliable. Swadesh assumed that basic vocabulary is replaced at a constant rate, but this assumption no longer has wide acceptance. A language that has undergone more rapid change than its sisters will appear more distantly related to them than it really is.

In her study of worldwide typological diversity, Nichols (1992: 24–25) adopted the units 'family' and 'stock.' The family she defined as a group with about the time depth of one of the older branches of Indo-European (2500–4000 years, e.g., Iranian, Balto-Slavic), recognizable by inspection when regular correspondences between word forms

and morpheme paradigms are displayed. The stock is the deepest phylogenetic node at which a protolanguage is reconstructable by the comparative method (5000–8000 years, e.g., Indo-European, Austronesian). Nichols (1997) adds the 'quasi-stock,' a grouping of stocks with promising phylogenetic markers but with no regular sound correspondences and few clear cognates. By this rubric, Afro-Asiatic, for example, is a quasi-stock. This approach lacks quantitative support, but it has the advantage that groups of languages under comparison meet the same methodological requirement.

The failure of quantitatively defined language groupings has a further consequence. The quantitative approach assumed languages to be composed of dialects, and families to be composed of languages. Since there is no quantitative distinction between a dialect and a language, linguists sometimes fall back on the assertion that dialects are mutually comprehensible, languages not. But as mutual comprehensibility is also a matter of degree, the distinction has limited objective validity, and 'lect' will be used here as a convenient cover term for both (*see* **Language and Dialect: Linguistic Varieties**). It follows that if there is no absolute distinction between the dialects of a language and the languages of a family, there is similarly no absolute distinction between a language with dialects and a family.

Complex Internal Structures

The tree structure of some large families, e.g., Indo-European and Austronesian, has been worked out in some detail. The structures of others have not. There are at least three reasons for this. Some putative families such as Afro-Asiatic are generally accepted, even though their time depth is apparently greater than 8000 years. The data reflect too much change and too much divergence to allow a thorough reconstruction of the protolanguage. Other families, such as Sino-Tibetan, have a history of frequent migrations back and forth, in the course of which lects have diversified but remained in contact, influencing each other in ways that make it very difficult to sort out borrowing from shared inheritance (LaPolla, 2001). A third reason is that the relationships among languages within many families are very complicated. In many – and perhaps most – parts of the world, the idealized model of a language family as the outcome of a community of speakers dividing into discrete daughter communities does not fit the data. Instead, a family may arise through the differentiation of an expanding community's speech into a network of lects, so that speakers of most lects can understand those of communities within a certain radius, but comprehension diminishes the further the speaker moves away from home.

Language families that arise in this way are the subject of Johannes Schmidt's (*q.v.*) wave theory, whereby innovations spread out from the center of a network like the ripples when a stone is thrown into a pond. The tree and wave models resist integration into a single model, and this has often been taken as a sign that they are in conflict with each other. This is unfortunate, as they model different phenomena. In the tree model, a community splits, and the lect of one or more of the new communities undergoes innovations that are subsequently inherited into its daughter lects. In the wave model, a community spreads, and innovations spread at different rates through the resulting network. The tree model fits large amounts of Austronesian and Uralic data rather well, while the wave model works better in other areas. However, we would like to be able to examine the history of a language family within a single framework, and the discussion below examines how lectal differentiation can be interpreted within a tree model and how the latter needs to be modified to accommodate it.

Dutch and German taken together provide a testbed. They form a network of lects covering the Netherlands and Germany (except the Friesian Islands), nearly half of Belgium, Luxembourg, two-thirds of Switzerland, Liechtenstein, and most of Austria. Until recently, this continuum extended into Alasace and into areas in northern Italy. If we for a moment ignore speakers' present-day (and often fairly recent) ability to communicate with each other in varieties of either standard Dutch or standard German, then we have a situation in which everyone readily understands nearby lects and in which there are few major boundaries between groups of lects, yet lects at greater distances from each other, especially on a north–south axis, are mutually incomprehensible. How does one split such a network up into languages in order to represent it in a family tree? Not surprisingly, attempts to do so disagree. Whatever higher-order groups are posited, each includes lects that are geographically contiguous with resemblant, mutually comprehensible lects in other higher-order groups, reflecting historical relationships that the grouping belies. A tree cannot do justice to the historical relationships. Instead, the best we can do is to redraw Figure 2, which shows the West Germanic part of Figure 1, as Figure 3. However, this avoids too much distortion by representing dialect networks very grossly. Norwegian, Danish, and Swedish similarly form a dialect network, and so does English in Britain and Ireland.

The presenting problem with the Dutch–German network is that its speakers (and most linguists)

would call its lects dialects, yet the degree of diversity across the network as a whole is intuitively greater than we expect in a single language. In other words, the network is like a (small) family made up of dialects, with no intervening level of 'language.' This problem disappears if one recognizes that the distinction between 'dialects' and 'languages' is artificial and a matter of degree of divergence. But it leaves unanswered the question of whether the Dutch–German network is more appropriately called a language or a family.

This question can be answered from two perspectives. One emanates from the line of thought above: the difference between language and family is also one of degree of diversity, and the question is an artifact of an unsubstantiated terminological distinction. But this is a little too simple if we are using 'family' recursively, because, ascending the tree, there comes a point at which the diversity within a collection of lects requires us to call it a family.

It might at a pinch be reasonable to call the Dutch–German network a language, but it would be unreasonable to call West Germanic, the next node up the tree, anything but a family, since English, Friesian, Afrikaans, Dutch–German, and Yiddish are all quite distinct from each other. Where does this distinctness reside? Partly in their degree of difference from each other and partly in the fact that they are not linked by transitional dialects. It is tempting to attribute these facts to geography, as the languages are mostly not contiguous, but the crucial factor is the strength of the social boundaries that separate speech traditions. For example, Afrikaans is spoken in South Africa, far from the Dutch lects of which it is an offshoot. Although they remain largely mutually comprehensible, their differences are striking. But if geography were the major factor, we should expect the Englishes of Britain/Ireland and North America to have diverged to a similar degree, yet they have not. After the establishment of British rule in 1795, speakers of the divergent southern African dialect of Dutch became socially isolated from their fellow-speakers in the Netherlands and a separate standard emerged.

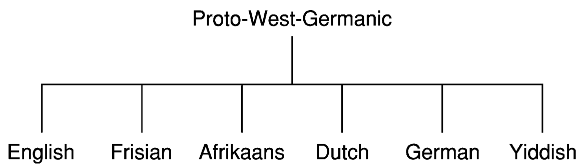


Figure 2 A conventional family tree of West Germanic.

Looking at the Dutch–German network from the perspective of social boundaries, we find that German speakers recognize dialect groupings such as Low German, Swabian, and Swiss German. Despite fuzzy boundaries, these groupings also have a reality for dialectologists, but no one would normally call them languages. One reason for this is that their social boundaries are also fuzzy, even across national borders. In this context the application of the term 'language' is as much a product of political as of linguistic history. The local lects of speakers who also speak a variety of standard German (in Germany, Switzerland, Liechtenstein, and Austria) are considered to be German dialects. The local lects of those who also speak a variety of standard Dutch (in the Netherlands and Belgium) are considered to be Dutch dialects. The linguistic repertoire of today's dialect speakers also includes their variety of the standard. In Switzerland, speakers are often diglossic in a Swiss German dialect and Swiss-flavored standard German, i.e., there is a measure of separateness between the two systems. Elsewhere, a speaker's repertoire is a continuum from a local dialect to a version of the standard, and s/he moves back and forth along the continuum according to whom s/he is speaking with. In these speakers, dialect and standard usually influence each other, with the result that the mutual comprehensibility of local lects across the Dutch–German border is being reduced as speakers accommodate to standard Dutch or standard German, and the standard languages are thereby reinforcing the social boundaries between the political entities with which they are associated.

The non-absoluteness of dialect/language and language/family distinctions is also manifest in Sinitic studies. The Sinitic network has conventionally been described as the Chinese language, comprising the Chinese dialects, but the Chinese language is comparable in diversity to the West Germanic family. This terminological situation has arisen because the Chinese language has long been coterminous with the political and social entity of China. The Chinese dialects and Chinese language are now sometimes

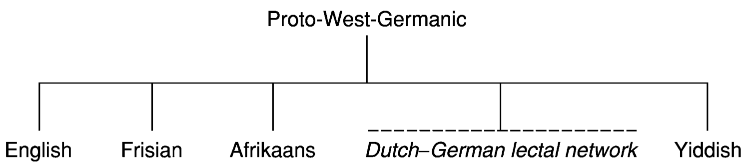


Figure 3 A revised family tree of West Germanic.

called the Sinitic languages and the Sinitic family, leaving the term 'Chinese language' to denote the standard language.

Relationships between dialect and standard and relationships among dialects are not easily rendered in a family tree diagram. Nor are situations where the historically closest sisters of a language, e.g., Afrikaans, are themselves certain dialects of another language, Dutch.

Extensive lectal networks occur in many parts of the world. New Guinea and Island Melanesia host a number, although none with the geographic extent of Dutch–German. Because their speakers' repertoire usually does not include a related standard language, the dialect/language question is not raised, but the absence of standard languages over millennia has permitted quite complex linguistic events when social boundaries have shifted. The complex history reconstructed in Figure 4 shows one area where networks have broken and later rejoined in a different configuration as the result of shifting boundaries. The basis for this reconstruction is the distribution of innovations. The Austronesian lects of New Ireland, a large island to the east of New Guinea, share certain innovations and appear at first sight to reflect the differentiation of a single lect, Proto-New-Ireland, into a lectal chain running the length of the island, then into a series of small families. But the lects of the southernmost family share other innovations with lects spoken on islands to the southeast. First, it seems, a single speech community rapidly settled New Ireland, and their speech differentiated into a chain of lects. Initially, population density was low, and social boundaries emerged between lects. The southernmost lect underwent certain innovations, and then some of its speakers sailed east and settled Nissan Island. From Nissan they established settlements on Buka Island to the south. The result was the South New Ireland/Northwest Solomonic network. Initially, frequent contact was maintained across the new network, more frequent than with other communities in New Ireland. Later on, populations increased, and the south New Ireland community resumed contact with communities to its north and was reintegrated into a New Ireland lectal network, through which innovations now passed (i.e., limited koineization occurred). Links with Nissan diminished to an annual voyage. The speech of the northwest Solomons community underwent certain innovations and became Proto-Northwest-Solomonic, then its speakers scattered through parts of Bougainville and the New Georgia group and across Choiseul and Santa Isabel to form the communities where the member lects of today's Northwest Solomonic family are spoken (Ross, 1988: 216–218, 258–259, 293–313).

Although the lects of New Ireland look as if they form a single node in the tree, the analysis of shared innovations reveals a much more complex history. A similar but more complex series of events has occurred in the history of the Fijian lects and the genesis of Proto-Polynesian (Geraghty, 1983, summarized by Ross, 1997: 227–229).

Cases such as New Ireland and Fiji raise a question of method. If, as in southern New Ireland, two sets of innovations overlap in a language or a group of languages, then they cannot both reflect a shared ancestor. One or both sets must reflect diffusion across lectal boundaries. How does one determine whether one set (or neither) reflects a shared ancestor? The answer lies in the nature of the innovations. The innovations common to the languages of southern New Ireland and the northwest Solomons entail bound morphology, which is only very rarely subject to diffusion. The innovations common to New Ireland as a whole do not, and are candidates for diffusion.

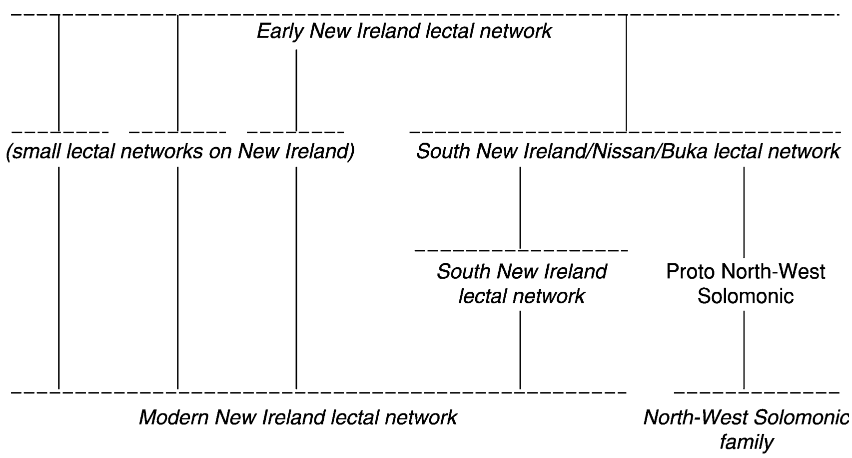


Figure 4 Network breaking and joining in New Ireland (after Ross, 1997: 230).

Mechanisms of Diversification

A family tree diagram can lull one into the false assumption that diversity occurs between lects, not within them, and that lects at terminal nodes are homogeneous. However, we know that an individual's speech varies – for example, along a continuum from local lect to standard – and that the degree to which variables are manifested differs among speakers. A speaker's choice of variables on particular occasions is in a complex relationship to her personal social network. We can distinguish between primary and secondary networks (Nettle, 1999: 67). At least in traditional societies, a person's primary network is likely to be dense, multiplex, kin-based, and enduring, whereas links in the secondary network have single functions such as trade. This distinction is a simplification, but a helpful one. Grossly, the primary network determines the features of a person's speech, while the secondary network mediates changes in those features.

Social network research shows that speakers are likely to use variables – pronunciations, words and phrases, grammatical constructions – which are used by and identify them with others in their primary network, although age and gender are also determinants (Milroy, 1987, 2001). The choice of variables on a particular occasion is also biased by the nature of that occasion and by the speech of the interlocutor. There is thus some diversity in even the smallest unit of a network, namely a speaker's primary links.

How do innovations enter a tight-knit primary network? The answer seems to be that its members typically have weak ties with outsiders through their secondary network, who in turn, because of their social roles, have weak ties with members of other primary networks, and it is these multiple-weak-tie individuals who are less subject to the norm enforcement of a primary network and act as carriers of innovation across the larger network (Milroy and Milroy, 1985).

How an innovation begins is difficult to investigate, but it seems that a speaker repeats a variant (random or deliberate) which is acquired by the speaker's children and/or copied by other speakers and selected as the marker of a social group (Weinreich *et al.*, 1968; Milroy, 1992: 170; McMahon, 1994: 251; Croft, 2000: 44–78). Among the Takia of Karkar Island (Papua New Guinea), there is a division between coastals and inlanders. Coastals distinguish the phonemes /l/ and /r/. Inlanders merge them as /l/. Every Takia adult knows about this difference. Comparative evidence shows that the innovators are the inlanders. The merger of /l/ and /r/ as inland /l/ must have occurred randomly in the speech of a single speaker, been acquired by others, selected as an inlander marker, and carried from one primary network to another by inlanders with multiple secondary links. Nettle (1999: Chaps. 2–3) finds that an innovation is unlikely to catch on without the amplification afforded by social selection.

When a speech community becomes two or three communities and contact among the new communities is reduced, their speech may diverge and new lects appear. The divergent changes are driven by social selection. The deciding factor behind divergence is the strength of social boundaries – the weakening of links between social networks so that new social identities emerge.

Some languages are far more resistant to change than others. This difference has been attributed to social network structure. Milroy and Milroy (1985) propose that when the speakers of a language form a set of overlapping tight primary networks, i.e., a tight-knit speech community, with only weak social links to other groups, the language typically changes very slowly. The paradigm case is Icelandic, little changed since early medieval Old Norse, unlike its radically transformed sisters Norwegian, Danish, and Swedish. At the opposite extreme are papuanized Austronesian languages in mainland New Guinea. Their speakers entered into symbiotic relationships with speakers of Papuan languages so that their secondary links were with Papuan speakers whose alien speech was the source of extensive innovations (Ross, 1996). However, Labov (2001: Chap. 10) offers an occupation-based interpretation of data sets like the Milroys', and there are relatively isolated communities (e.g., in Polynesia) where change has been quite fast, and apparently even tight-knit communities where speakers have exaggerated differences from their neighbors, i.e., accelerating change to maintain isolation (Anderson, 1988; Thurston, 1989; Ross, 2003).

Nettle (1999: 66–78) argues that economic (inter) dependence is the main reason for language spread. He contrasts the many small speech communities of inland New Guinea, each of them economically largely independent, with the widespread communities of the Hausa, the Fulani, and the Tuareg of on the southern edge of the Sahara, where survival requires distant economically based relationships. Hunter-gatherer pygmy groups in central Africa depend on their farmer neighbors to supplement their diet, and each group of pygmies has shifted to the language of the farming group with which it is paired. Hunter–gatherer negrito groups in the Philippines have adopted Austronesian languages from farmer neighbors (Reid, 1994). The difference between these situations and the one which led to the papua-

nization of Austronesian languages in New Guinea was probably a difference in the strength of primary network links: these links were stronger among Austronesian-speaking horticulturalists than among hunter- gatherers.

With explicable exceptions, language density is greater nearer the Equator than further away from it, perhaps because it is correlated with ecological risk: the more difficult it was to sustain human life, the larger the economic and therefore linguistic networks that came into being (Nettle, 1999: 60–66, 79–93). This economy–language pairing continues in more recent events. The expansion of European colonizers across the planet since the 15th century caused a devastating reduction in language diversity, now being intensified by the push toward economic globalization.

The Origins of Phylogenetic Diversity

The origins of the phylogenetic diversity of languages lie far back beyond knowability. Are all today's language families descended from a single 'Proto-World'? Many people have assumed so, not on linguistic evidence but on the basis of a single human origin in east Africa. The comparative method allows us to reconstruct linguistic history only as far back as about 8000 years, yet structurally modern language has probably been around for 100 000 years.

Nichols (1992: 221–230) finds a geographic patterning of morphosyntactic features in the world's languages, which, she suggests, are fossil reflexes of the original spread of languages. She traces a path from the Old World into the Pacific and rapidly south to Australia, followed by circum-Pacific and New World colonization. Those who entered the Americas across the Bering Strait were related to a circum-Pacific group, not an Old World population.

Using Nichols' stocks, Nettle (1999: 113–129) divides the world into nine regions and plots stock and language densities for each of them. Setting aside the New Guinea figures, which are significantly higher than anywhere else, he finds that stock (phylogenetic) density in Africa, Europe, and Asia is far lower than in the rest of the world. In the rest of the world there is a simple correlation: the more languages, the more stocks. This situation is claimed to reflect time depth. Over time, interaction between human populations has led through language shift to stock extinctions. On archaeological estimates, the time depth of human language in Africa is more than 100 000 years; it ranges from 60 000 to 40 000 years in Europe, Asia, Australia, and New Guinea; and is as little as 12 000 years in the Americas (although this dating is controversially recent). Nettle's stock densities are in an approximate inverse relationship to these time depths.

Bellwood (1997) argues that the Neolithic transition, i.e., the transition from foraging to agriculture, caused the expansion of some of the world's largest language families, presumably leading to increased extinctions of languages and stocks. He notes four significant cultivation events that apparently caused the expansion of languages into often large families, probably replacing their own closest relatives as well as other families in the process:

1. Wheat, barley and legumes in the Fertile Crescent by 8000 BC: Afro-Asiatic, (controversially) Indo-European, Elamo-Dravidian, Kartvelian.
2. Taro and bananas in the highlands of New Guinea by 8000 BC (Denham *et al.*, 2003): Trans-New Guinea family, limited to the New Guinea region.
3. Rice, foxtail and broomcron millet in the Yangtze and Yellow River Basins by 6000 BC: Austronesian, Tai-Kadai, Austro-Asiatic, Sino-Tibetan, Hmong-Mien.
4. Sorghum and pearl millet in sub-Saharan Africa by 2000 BC (?): Nilo-Saharan and Niger-Congo.

These events may have sharply reduced the world's phylogenetic diversity. Relics of what was there before perhaps survive in Basque in the Pyrenees, in the diversity of the language families of the Caucasus, and of parts of New Guinea and parts of the Americas (see Nichols, 1997).

In Diamond and Bellwood (2003), the Neolithic-transition hypothesis is extended across the world, but somewhat tentatively, and wisely so. Although it is probably correct for Austronesian, and no doubt elsewhere, it would be wrong to overgeneralize it. Nilo-Saharan and Niger-Congo are not necessarily phylogenetic units (Nichols, 1997). Pama-Nyungan in Australia, Uralic, and Chukchi-Kamchatkan in northern and central Eurasia, Khoisan in southern Africa, and Athabaskan and Eskimo-Aleut in northwest North America have all expanded without a Neolithic transition, although there may have been other economic reasons for their spreads. Various agricultural families have not expanded significantly: Ramu–Lower Sepik in New Guinea, Nakh-Daghestanian and Kartvelian in the Caucasus, and various families in the Americas (Campbell, 2002). The putative Indo-European agricultural expansion (Renfrew, 1987) has been called into question by Nichols (1998), who brings a cohort of arguments, linguistic, geographic and historical, for a homeland in western Central Asia around 3500 BC. This is part of a larger argument that stock densities are also heavily influenced by geographic factors (Nichols,

1997), to which Campbell (2002) would add differences in social behaviors among speech communities.

A New Evolutionary Model

Recently, an evolutionary model has been proposed in which the replicator is what Croft (2000) calls the lingueme. Its basis is a general theory of evolution applicable to change in social phenomena as well as in biology. There are three versions of the model (space compels us to ignore their differences): Nettle (1999), Croft (2000), and Enfield (2003). Croft's is the most articulated version. A lingueme is any unit of linguistic structure, be it a phoneme, a morpheme, a word or phrase, or a construction, i.e., the units which are the parameters of intra- and interspeaker variation. The linguistic counterpart of the biologist's DNA string is the utterance, a structured set of linguemes. The speaker, as the repository of grammar, corresponds to the biological organism (the interactor). As in network theory, speakers form a networked population, and language change occurs when a lingueme is propagated across the network in altered form and becomes manifest in speakers' utterances. Croft (2000: Chaps. 5–6) builds in a paradigm of possible innovation types as well.

What is new and what gives the model its integrative power is the view of the lingueme as the replicator. The language is deliberately left out of focus, recognizing that a speaker's repertoire is a structured collection of variables that do not necessarily have common geographical boundaries. The dialect vs. language vs family issue does not arise.

What is somewhat uncertain, however, is where this leaves the grammar. Enfield stresses that language learners infer "grammatical patterns" from others' behavior. The idea that a single grammar is shared by speakers or is transmitted from one generation to the next is for him a cultural illusion. Linguistic signs "are best understood as theories, constructed by individual speakers over time by a process of trial and error" (Enfield, 2003: 2–3). Sets of signs come to cohere as systems in individuals' minds, and speakers' grammars have a lot in common because of the need for coordination. Croft (2001: 29) grants that the evolutionary model implies looser grammatical organization than either the structuralist or the generative model, but attributes "a high degree of structural organisation" to the lingueme pool. He asserts that through the utterances she hears, a child inductively acquires syntactic constructions, and then from these infers a taxonomic network of constructions (Croft, 2001: 57–58).

Nettle, Croft, and Enfield all give pride of place to contact phenomena as a justification for the new model. They point out that if the replicator is the lingueme, then it does not matter whether an altered lingueme has its origin in what speakers recognize as "their own language" or "another language." The process of change is the same. Under the family tree model, a language is only allowed to have one parent, and contact phenomena, they imply, tend to be pushed under the carpet. This is true, but the emphasis on contact risks overlooking what is modeled in a family tree, namely the generational continuity of a language. Although grammar is not itself a replicator and is reconstructed by each new speaker, parents and children usually have a clear sense that they are speaking the same language. Furthermore, contact in most cases entails bilingualism, which causes the linguemes of one language to be modified under the influence of the other. The modified language is usually still recognized as the same language by the next generation, whether altered linguemes have a language-external or language-internal origin. This is so even for the papuanized Austronesian languages of New Guinea. The generational continuity depicted by a family tree is only broken when a social catastrophe occurs, as when the transportation of Melanesians to far-off plantations from the 1860s led to the rapid stabilization of Pacific Pidgin, a language with no previous existence as a system (Ross, 1997: 251–253). Language contact may have been marginalized by the family tree model, but it would be a pity if generational continuity were marginalized under the new model.

See also: Cladistics; Contact-Induced Convergence: Typology and Areality; Cultural Evolution of Language; Early Historical and Comparative Studies; Evolutionary Theories of Language: Current Theories; Evolutionary Theories of Language: Previous Theories; Fijian; Genetics and Language; Germanic Languages; Labov, William (b. 1927); Language and Dialect: Linguistic Varieties; Language Change and Language Contact; Language/Dialect Contact; Long-Range Comparison: Methodological Disputes; Microparametric Variation; Origin and Evolution of Language; Papua New Guinea: Language Situation; Phonemics, Taxonomic; Solomon Islands: Language Situation; Subgrouping Methodology; Variation and Language: Overview; Variation in German.

Bibliography

Andersen Henning (1988). 'Centre and periphery: adoption, diffusion and spread.' In Fisiak J (ed.) *Historical dialectology.* Berlin: Mouton de Gruyter. 39–85.

Bellwood P (1997). 'Prehistoric cultural explanations for the existence of widespread language families.' In McConvell P & Evans N (eds.) *Archaeology and*

linguistics: Aboriginal Australia in global perspective. Melbourne: Oxford University Press. 123–134.
Campbell L (2002). 'What drives linguistic diversification and language spread?' In Bellwood P & Renfrew C (eds.) *Examining the farming/language dispersal hypothesis.* Cambridge: McDonald Institute of Archaeological Research. 49–63.
Croft W (2000). *Explaining language change: an evolutionary approach.* Harlow: Pearson Education.
Croft W (2001). *Radical construction grammar: syntactic theory in typological perspective.* Oxford: Oxford University Press.
Denham T P, Haberle S G, Lentfer C, Fullagar T, Field J, Therin M, Porch N & Winsborough B (2003). 'Origins of agriculture at Kuk Swamp in the Highlands of New Guinea.' *Science 201,* 189–193.
Diamond J & Bellwood P (2003). 'Farmers and their languages: the first expansions.' *Science 300,* 597–603.
Enfield N J (2003). *Linguistic epidemiology: semantics and grammar of language in mainland Southeast Asia.* London: RoutledgeCurzon.
Geraghty P (1983). *The history of the Fijian languages.* Oceanic Linguistics special publication No. 19. Honolulu: University of Hawaii Press.
Labov W (2001). *Principles of linguistic change. 2: Social factors.* Oxford: Blackwell.
LaPolla R J (2001). 'The role of migration and language in the development of the Sino-Tibetan languages.' In Aikhenvald A & Dixon R M W (eds.) *Areal diffusion and genetic inheritance: problems in comparative linguistics.* Oxford: Oxford University Press. 225–254.
McMahon A (1994). *Understanding language change.* Cambridge: Cambridge University Press.
Milroy J (1992). *Linguistic variation and change: on the historical sociolinguistics of English.* Oxford: Blackwell.
Milroy J & Milroy L (1985). 'Linguistic change, social network and speaker innovation.' *Journal of Linguistics 21,* 339–384.
Milroy L (1987). *Language and social networks* (2nd edn.). Oxford: Blackwell.
Milroy L (2001). 'Social networks.' In Chambers J K (ed.) *The handbook of language variation and change.* Oxford: Blackwell. 549–572.
Nettle D (1999). *Linguistic diversity.* New York: Oxford University Press.
Nichols J (1992). *Linguistic diversity in space and time.* Chicago: Chicago University Press.
Nichols J (1996). 'The comparative method as heuristic.' In Durie M & Ross M D (eds.) *The comparative method reviewed: irregularity and regularity in linguistic change.* New York: Oxford University Press. 39–71.
Nichols J (1997). 'Modeling ancient population structures and movement in linguistics.' *Annual Review of Anthropology 26,* 359–384.
Nichols J (1998). 'The Eurasian spread zone and the Indo-European dispersal.' In Blench R M & Spriggs M (eds.) *Archaeology and language, 2: Correlating archaeological and linguistic hypotheses.* London: Routledge. 220–266.
Reid L A (1994). 'Unravelling the linguistic histories of Philippine negritos.' In Dutton T E & Tryon D T (eds.) *Language contact and change in the Austronesian world.* Berlin: Mouton de Gruyter. 443–475.
Ross M D (1988). *Proto Oceanic and the Austronesian languages of western Melanesia.* Canberra: Pacific Linguistics.
Ross M D (1996). 'Contact-induced change and the comparative method: cases from Papua New Guinea.' In Durie M & Ross M D (eds.) *The comparative method reviewed: regularity and irregularity in language change.* New York: Oxford University Press. 180–217.
Ross M D (1997). 'Social networks and kinds of speech-community event.' In Blench R M & Spriggs M (eds.) *Archaeology and language, 1: Theoretical and methodological orientations.* London: Routledge. 209–261.
Ross M D (2003). 'Diagnosing prehistoric language contact.' In Hickey R (ed.) *Motives for language change.* Cambridge: Cambridge University Press. 174–198.
Ross M D (2005). 'Pronouns as a preliminary diagnostic for grouping Papuan languages.' In Pawley A, Attenborough R, Hide R & Golson J (eds.) *Papuan pasts: Investigations into the cultural, linguistic and biological history of the Papuan-speaking peoples.* Canberra: Pacific Linguistics.
Swadesh M (1972). *The origin and diversification of language.* London: Routledge and Kegan Paul.
Thurston W R (1989). 'How exoteric languages build a lexicon: esoterogeny in West New Britain.' In Harlow R & Hooper R (eds.) *VICAL 1, Oceanic languages: papers from the Fifth International Conference on Austronesian Linguistics.* Auckland: Linguistic Society of New Zealand. 555–579.
Weinreich U, Labov W & Herzog M (1968). 'Empirical foundations for a theory of language change.' In Lehmann W P & Malkiel Y (eds.) *Directions for historical linguistics.* Austin: University of Texas Press. 95–195.

Language Identification, Automatic

P Juola, Duquesne University, Pittsburgh, PA, USA

Language identification is one of the most fundamental aspects of text analysis and processing. Before indexing a web page – or doing almost anything else – the analyst must first identify in which language the page is written. With the recent explosion in the number of written documents demanding analysis via the Web, this question has received recent research interest.

At least in theory, it is not a difficult problem. Cryptographers have known for centuries (Kahn, 1967) that some letters are more common than others, and that distribution varies between languages. A cryptographer usually starts by assuming (or guessing at) the language – an assumption to be abandoned if a solution is not forthcoming. For example, one could note that the letter 'W', although relatively uncommon in English, is almost nonexistent in French, and therefore a document containing a noticeable number of 'W's is probably not in French. Alternatively, the letter sequence ETAOINSHRDLU, representing the 12 most common letters in descending frequency (for English), has become a cliche, but it is still useful as a informal description. Any document that closely matches this sequence is probably in English, whereas any document that varies too greatly from this is probably not. This observation can be formalized using a statistical model such as a multinomial distribution over an independent letter set.

Beesley (1988), one of the first researchers in this area, identified three main 'translator approaches' used by humans without computer assistance to identify languages. One strategy lists the letters (particularly accented ones) used in a given language. A second lists common/short words (such as function words), whereas the third lists typical word endings. Beesley applied the computer to extend these lists to statistical distributions. Specifically, he compiled distributions of bigram (two-letter pairs) frequency for each of a set of languages. For a document in an unknown language, each component bigram is looked up in the appropriate distribution for each language of interest, and the probabilities are multiplied together. The language table yielding the highest overall probability is inferred to be the (unknown) language of the document. Although this technique is philosophically weak (it assumes, e.g., that document bigrams are independently distributed, whereas the effects of context on language are well-known), it still yields accurate results with a minimum of computation.

This method also lends itself well to generalizations: for example, calculating trigram (three-letter sequences) or, more generally, n-gram probabilities. The analysis can also be performed on word frequencies (or n-grams). Analysis can focus on specific words or letters as shibboleths, morphological characteristics, or even syllable characteristics. Similarly, the exact analytic framework used can vary, from the simple independent event statistics described above through dissimilarity measurements, discriminan analysis, crossentropy (or other) distances, Bayesian statistics, support vector machines, neural networks, or Markov processes (to name a few methods that have been applied).

From a theoretical perspective, language identification can be regarded as a specific instance of text categorization (Manning and Schütze, 1999), under which broad heading can be clustered other problems such as genre identification, authorship attribution, document dating, spam detection, or even in some cases theme and tone analysis. For this reason, many techniques developed for one problem (such as the techniques listed for authorship attribution in Holmes [1994]) can be applied in other areas, as shown by the work of Juola (1997, 2003). This raises the as-yet unsolved theoretical question about what kinds of information and categorization techniques are most appropriate for making the distinctions relevant for one particular subproblem.

In some regards, language identification from text is a solved problem, as researchers routinely cite successful classification rates of 99% or more. However, there are still practical considerations to be addressed. This extreme accuracy often requires 100 or more characters of test text: a small number, but still more than, for example, many newspaper headlines or cryptograms, and substantially more than a skilled human requires to make the same judgment. These methods can also have difficulty with segmentation, identifying those passages in a second language that add a certain *je ne sais quoi* to an otherwise monolingual document. These interloping phrases distort statistics and may confuse otherwise accurate methods. (Consider, e.g., the analysis of a typical legal brief, with the normal amount of 'dog-Latin' used in legal terms of art. It should be noted that the segmentation issue is not confined to language identification, as co-authorship or co-editing can pose similar problems for other text categorization tasks.) A more subtle problem with the approaches outlined is their demand for a suitable set of training data for each language under consideration; a simple implementation of a system will not usually respond 'None of the

above'. A system trained on English, German, French, and Spanish, for example, would probably identify a document in Dutch or Norwegian as being in German.

Finally, the method should be robust enough to handle realistic variation in document coding or typography. For example, the letters 'A' and 'a' are often considered identical, and some kinds of English text (e.g., newspaper headlines) are written using upper-case only. Similarly, many texts are written without using diacritical marks or similar language-specific features, often because of the limitations of the system on which they were written (e.g., a lack of eight-bit support on many text editors). More generally, the encoding of non-Latin letters (including Cyrillic and Greek) is often system dependent, and even the encoding of Latin letters with diacritical marks can be system dependent. Beesley cites as an example the decimal value 123, which can identically represent (in different 'standard' forms of EBCDIC) the octothorpe ('#') in English, 'Ñ' in Spanish, 'Æ' in Danish, 'Ã' in Continental Portuguese, and 'Õ' in Brazilian Portuguese. More seriously, the same orthographic character can often be mapped (depending on the system and processor) to different characters; the capital Œ ligature is character 0x8C in Microsoft's Code Page 1252 standard, but character 0x152 in Unicode (or standard HTML). A system cannot necessarily be expected to know the encoding used for a given document.

A similar problem occurs in the treatment of typographic or other errors. The digraph 'aa' is very rare (frequency zero in many training samples) in English but is easily produced by an 'aaccidental' doubling. A system that calculates the overall probability by multiplying individual digraph probabilities together would thus find the probability of this article (or this encyclopedia) being in English to be zero, on the basis of the forbidden digraph. With small enough test samples, this sort of error-introduced statistical distortion can be substantial.

An ideal method should be efficient enough for real-time processing, not make undue demands on system storage, be easily extensible to new languages, and be robust against the issues described above. Although such a system does not yet exist, language identification remains an active research area because of its fundamental importance as a base for other forms of automatic text processing.

As a final note, one interesting variation on language identification begins not with text but with images such as GIF files. Before an optical character recognition (OCR) system can work properly, it must apply a language model, based of course on the language in which the document is written. If this language is unknown, there is an impasse created. Researchers including Sibun, among others, have attempted to resolve this by using image features and basic character shape features, such as the presence or absence of descenders, that can be spotted using only image-processing techniques. These features can be used as events analogous to letters or words in the traditional (often *n*-gram-based) inference engines. Work has thus far largely been confined to Roman alphabets, but extension to other character sets and languages will no doubt continue, with interesting results. Automatic identification of language from handwritten documents might prove more interesting yet, but no research in this area has come to the author's attention.

At the most abstract level, there is a fundamental question that remains unresolved. Despite the accuracy obtainable by these simple statistical methods, researchers are no closer to answering questions about the fundamental differences between languages or, equivalently, how the low-level statistical features characteristically employed to the solution of language identification relate to the high-level aspects of language studied by traditional linguists. Image- and handwriting-based language identification (as well as the continuing development of identification systems for non-Latin alphabets) may be the next major challenge, in part because we have as yet no general notion of what, exactly, constitutes perceptual 'English'-ness (or other languages).

See also: Character Sets; Classification of Text, Automatic.

Bibliography

Beesley K R (1988). 'Language identifier: a computer program for automatic natural-language identification on on-line text.' In *Proceedings of the 29th Annual Conference of the American Translations Association.* 47–54.

Cavnar W B & Trenkle J M (1994). '*N*-gram-based text categorization.' In *Proceedings of the 3rd Annual Symposium on Document Analysis and Information Retrieval.* 161–175.

Damashek M (1995). 'Gauging similarity with *n*-grams: language independent categorization of text.' *Science 267,* 843–848.

Grefenstette G (1995). 'Comparing two language identification schemes.' In *Proceedings of 3rd International Conference on Statistical Analysis of Textual Data (JADT 95).* 263–268.

Holmes D I (1994). 'Authorship attribution.' *Computers and the Humanities 28,* 87–106.

Juola P (1997). 'What can we do with small corpora? Document categorization via cross-entropy.' In *Proceedings of*

the Workshop on Similarity and Categorization (SimCat 97).
Juola P (2003). 'The time course of language change.' *Computers and the Humanities 37(1),* 77–96.
Kahn D (1967). *The codebreakers.* New York: Scribner.
Manning C & Schütze H (1999). *Foundations of statistical natural language processing.* Cambridge, MA: MIT Press.
Muthusamy Y K & Spitz A L (1998). 'Automatic language identification.' In Cole R A *et al.* (eds.) *Survey of the state of the art in human language technology.* Cambridge: Cambridge University Press.
Sibun P & Reynar J (1996). 'Language identification: examining the issues.' In *Proceedings of the 4th Annual Symposium on Document Analysis and Information Retrieval.* 125–135.
Sibun P & Spitz A L (1994). 'Language determination: natural language processing from scanned document images.' In *Proceedings of the Fourth ACL Conference on Applied Natural Language Processing.* 115–121.
Teytaud O & Jalam R (2001). 'Kernel-based text categorization.' In *Proceedings of the International Joint Conference on Neural Networks.*

Language Ideology

J Blommaert, Institute of Education, University of London, UK, and Ghent University, Ghent, Belgium

Ideology and Language

Language ideology emerged as a separate field of linguistic-anthropological study in the last decades of the 20th century, combining linguistic ethnography with insights from the social-scientific study of ideology. Though the field is still very much under construction, its influence on linguistic anthropology, linguistics, discourse analysis, and sociolinguistics is considerable (Woolard, 1998).

Yet few social-scientific terms have had such complex histories of interpretation as the term 'ideology.' Nested in scholarship as well as in politics, and within scholarship in such widely diverging traditions as history, philosophy, political economy, anthropology, and linguistics, ideology has built a track record of controversy, dispute, and conflict over its meaning. As a social-scientific concept, it has at least two ancestors.

One tradition of conceptualizing ideology grew out of Marxist theory and came to identify the mediating link between material and ideational aspects of reality. In *The German ideology*, Marx and Engels presented ideology as 'false consciousness': systemically distorted perceptions of reality, more specifically of one's class position. Gramsci's *Prison notebooks* inserted ideology in the concrete struggle for political dominance and identified 'hegemony' as dominance by a particular class-bound ideology in the cultural and ideational-political field. Revolution, according to Gramsci, needed to attain hegemony as well as power in order to be successful. Later theorists such as Althusser carried the analysis of ideology further into a general sociology of power. According to Althusser, ideology was promulgated and sustained by "ideological state apparatuses": civil society institutions which 'normalized' the imaginary relationship between the subject and his/her position in the social world. To Habermas, ideology was lodged in communication distorted by power processes, and in Bourdieu's work, the conversion of hegemonic ideologies into 'normal' practices and ideas was labeled 'habitus.' In each of these developments, the Marxian conceptualization of ideology as tied to the interests of particular social groups and to processes of power and dominance was adopted and elaborated.

A second ancestor of ideology in the social sciences is the Durkheimian tradition of sociology as collective psychology. Durkheim's influence is noticeable in Marc Bloch's work on historical 'mentalities' as well as in a variety of anthropological approaches to 'world views' and 'belief systems' (as, for example, in Evans-Pritchard's work). A parallel development runs from German neo-Kantianism through Boas and Whorf. These ideational (cultural) complexes are, unlike in the Marxian tradition, neutrally defined: they are often presented as the deeper layers of culture and society, the unspoken assumptions that, as some kind of 'social cement,' turn groups of people into communities, societies, and cultures. This concept of ideology is often called the total concept, for it suggests the acceptance of ideational-cultural complexes by every member of the community. The Marxian and Durkheimian traditions of study are more or less each other's contemporaries, and both flourished in the 20th century.

Language has never been out of focus in the study of ideology, and important reflections on language and ideology emerge early in the Marx-inspired works of Bakhtin and Voloshinov, as well as somewhat later in the work of, for example, Roland

the Workshop on Similarity and Categorization (SimCat 97).
Juola P (2003). 'The time course of language change.' *Computers and the Humanities 37(1),* 77–96.
Kahn D (1967). *The codebreakers.* New York: Scribner.
Manning C & Schütze H (1999). *Foundations of statistical natural language processing.* Cambridge, MA: MIT Press.
Muthusamy Y K & Spitz A L (1998). 'Automatic language identification.' In Cole R A *et al.* (eds.) *Survey of the state of the art in human language technology.* Cambridge: Cambridge University Press.
Sibun P & Reynar J (1996). 'Language identification: examining the issues.' In *Proceedings of the 4th Annual Symposium on Document Analysis and Information Retrieval.* 125–135.
Sibun P & Spitz A L (1994). 'Language determination: natural language processing from scanned document images.' In *Proceedings of the Fourth ACL Conference on Applied Natural Language Processing.* 115–121.
Teytaud O & Jalam R (2001). 'Kernel-based text categorization.' In *Proceedings of the International Joint Conference on Neural Networks.*

Language Ideology

J Blommaert, Institute of Education, University of London, UK, and Ghent University, Ghent, Belgium

Ideology and Language

Language ideology emerged as a separate field of linguistic-anthropological study in the last decades of the 20th century, combining linguistic ethnography with insights from the social-scientific study of ideology. Though the field is still very much under construction, its influence on linguistic anthropology, linguistics, discourse analysis, and sociolinguistics is considerable (Woolard, 1998).

Yet few social-scientific terms have had such complex histories of interpretation as the term 'ideology.' Nested in scholarship as well as in politics, and within scholarship in such widely diverging traditions as history, philosophy, political economy, anthropology, and linguistics, ideology has built a track record of controversy, dispute, and conflict over its meaning. As a social-scientific concept, it has at least two ancestors.

One tradition of conceptualizing ideology grew out of Marxist theory and came to identify the mediating link between material and ideational aspects of reality. In *The German ideology*, Marx and Engels presented ideology as 'false consciousness': systemically distorted perceptions of reality, more specifically of one's class position. Gramsci's *Prison notebooks* inserted ideology in the concrete struggle for political dominance and identified 'hegemony' as dominance by a particular class-bound ideology in the cultural and ideational-political field. Revolution, according to Gramsci, needed to attain hegemony as well as power in order to be successful. Later theorists such as Althusser carried the analysis of ideology further into a general sociology of power. According to Althusser, ideology was promulgated and sustained by "ideological state apparatuses": civil society institutions which 'normalized' the imaginary relationship between the subject and his/her position in the social world. To Habermas, ideology was lodged in communication distorted by power processes, and in Bourdieu's work, the conversion of hegemonic ideologies into 'normal' practices and ideas was labeled 'habitus.' In each of these developments, the Marxian conceptualization of ideology as tied to the interests of particular social groups and to processes of power and dominance was adopted and elaborated.

A second ancestor of ideology in the social sciences is the Durkheimian tradition of sociology as collective psychology. Durkheim's influence is noticeable in Marc Bloch's work on historical 'mentalities' as well as in a variety of anthropological approaches to 'world views' and 'belief systems' (as, for example, in Evans-Pritchard's work). A parallel development runs from German neo-Kantianism through Boas and Whorf. These ideational (cultural) complexes are, unlike in the Marxian tradition, neutrally defined: they are often presented as the deeper layers of culture and society, the unspoken assumptions that, as some kind of 'social cement,' turn groups of people into communities, societies, and cultures. This concept of ideology is often called the total concept, for it suggests the acceptance of ideational-cultural complexes by every member of the community. The Marxian and Durkheimian traditions of study are more or less each other's contemporaries, and both flourished in the 20th century.

Language has never been out of focus in the study of ideology, and important reflections on language and ideology emerge early in the Marx-inspired works of Bakhtin and Voloshinov, as well as somewhat later in the work of, for example, Roland

Barthes and Raymond Williams. For Bakhtin and Voloshinov, language embodies and articulates the experience of social struggle, transition, and contest, and consequently the linguistic sign is seen as deeply ideological. Language usage displays a variety of orientations to social interests, derived from particular positions in society ('voices'). And interaction involves different voices evaluatively responding to each other's statements ('dialogue'). Thus, human communication through language displays meaningful metalevel inscriptions, adding a layer of sociopolitical, ideological meaning to the event. The works of Bakhtin and Voloshinov became one building block for the contemporary study of language ideology.

Another building block was Peircian semiotics and its emphasis on the nonreferential aspects of meaning. Importantly, the sign in Peirce's semiotics stood in a close relationship to objects in the world, and Saussurean semiotics – restricted to referential relations between signs and signified – was extended to include signification. Peirce developed a typology of signs in which he distinguished between 'icons' (signs that resemble their object), 'indexes' (signs that bear an existential relationship with their object), and 'symbols' (signs that conventionally represent an object as a member of a class of like kinds). And such signs operate in an infinitely generative way, signs generating other signs through processes of implicature, entailment, and the like (Keane, 2003). In such a way, the iconic and indexical dimensions of signs become part of the context in and on which signs operate.

This complex body of theory is brought to bear on ethnography in the work of Michael Silverstein, providing the origin of language ideology as a separate field of study. Revisiting Whorf's work on linguistic categories, Silverstein claimed that the 'world views' identified by Whorf were in fact indexically organized language-ideological complexes, and that "the indexical plane of meaningfulness properly encompasses the folk realm of rhetoric (the system of language use), how language signals derive their socially understood effects in various socially constituted situations of discourse" (Silverstein, 1979: 205). Such folk realms, he further argued, displayed "the tendency to rationalize the pragmatic system of language, in native understanding, with an ideology of language that centers on reference-and-predication" – a referential ideology of language (1979: 208). Focusing on language ideologies would allow the analysis of language to be the analysis of culture (cf. Silverstein, 1976, 2004), for precisely the indexical level of language would anchor it firmly into culture. Furthermore, language and language use could be seen as reflecting on itself, for every act of communication (pragmatics) articulates a metapragmatics in which the linguistic-ideological features operate. This metapragmatic, indexical layer of semiotic systems such as language is not neutral; it is evaluative, relational, socially positioned, invested with interests, and subject to contestation and dominance.

Whereas Silverstein's approach to language ideology is firmly embedded in the ethnographic tradition of linguistic anthropology, a more or less simultaneously developing tradition of studies focused on linguistic metatheory as ideology (Joseph and Taylor, 1990) and on popular and institutional language ideologies (Milroy and Milroy, 1985; Cameron, 1995). The sources of this other tradition are more diffuse, ranging from the history of science to variationist sociolinguistics and applied linguistics. In what follows, the focus will be on the ethnographic tradition, with an occasional foray into other bodies of work.

Ideology and/in Analysis

One of the effects of addressing language ideology is the fact that it dislodges a range of established concepts and categories and thus offers infinite opportunities for revisiting existing scholarship. I will discuss the impact of language ideologies on the following topics of inquiry: (1) central linguistic notions such as 'language' and (2) 'text'; (3) central sociolinguistic concepts such as 'speech community,' with implications for the study of language policy and (4) language change; (5) the history of linguistics.

Language

There is a longstanding concern in ethnography to investigate specific varieties of language rather than language *per se*. The reason, equally old, is that a uniformizing, singularized notion of language obscures the crucial sociolinguistic differences that occur within that language, and "it is a fallacy to equate the resources of language with the resources of (all) users" (Hymes, 1996: 213). In fact, the existence of 'language' and 'languages' – objects that are countable and have a name, such as English, Zulu, or Japanese – is a powerful language-ideological effect, the result of long historical processes of construction and elaboration of a metaphysics of mind vs. world within which the Cartesian self as sign, with the world as its object, comes into view.

In order to understand such processes, we have to start from Silverstein's referential ideology of language. Given the widespread belief in denotation as the main function of language, language itself is conceptualized as a transparent, structured, and finite system of clear (denotational) and noncontextual forms which characterizes groups of people. This we

could call an 'artifactual' view of language, a view in which language is seen as a manipulable, bounded artifact consisting of (grammatical) 'structures' with a clear function, denotation. Such a view would be expressed, for instance, in utterances such as 'I need to work a bit on my French' or 'His German needs some polishing, it is a bit rusty,' in which 'a language (name)' is metaphorically seen as an object one can obtain, possess, manufacture, and improve upon. This view, as we shall see, is a key ingredient of modernity and thus a rather recent construct, but it has become the most widespread view of language both in popular and in scientific circles. Linguistics has contributed in no small degree to the cultural construction of language in general as a stable, contextless individual mental object, and language and educational policies as well as larger nation-building programs have been deeply influenced by this ideology.

The 'language' thus artifactually conceptualized as a denotational code characterizing groups of people always takes concrete shapes. It would be qualified as a 'standard' variety; it would be strongly tied to literacy and focused on common grammatical structure and lexicon; and it would be the object of normative control, of institutional regimentation and orientation towards centers of authority (Silverstein, 1996). The 'standard' is usually perceived as 'neutral' or 'unmarked,' i.e., something the qualities of which are perceived as natural, self-evident, 'normal.' Thus, 'standard' English would often be qualified as 'accentless' English – English which cannot be characterized as to class or regional belonging – whereas in fact, obviously, such a variety would be strongly accented and solidly indexical of social, educational, or even regional and generational backgrounds (Agha, 2003). Similarly, writing in the 'standard' orthography would be unmarked ('correct'), while writing in an erratic orthography would be 'wrong' and would trigger indexicalities of sloppiness, lack of education, and so forth. 'Standard' is a particular variety of language, a register perceived as 'neutral' because of elaborate sociohistorical processes of normalization and codification (Silverstein, 1996; Blommaert, 1999a; Kroskrity, 2000a; Gal and Woolard, 2001a; see also Bourdieu, 1991; Cameron, 1995). Such processes often involve the construction and reproduction of specific codes and genres for public purposes, creating a sense of authority emanating from a center, marking particular forms of speech as emblematic of group belonging and identity, and introducing a sociopolitical evaluative stratification in language usage, with 'better' and 'worse' forms of usage (Blommaert, 2004; cf. the essays in Kroskrity, 2000b and Gal and Woolard, 2001b). Members of that group are expected to perform these codes and genres as part of their 'culture of language use,' indexing 'full' group membership and categorical belonging (Silverstein, 1996: 291).

Dislodging the notion of 'language' and identifying it as a complex of metapragmatic qualifications projected onto situated language usage has far-reaching critical import, for it shifts attention away from stable, contextless linguistic notions to deeply socioculturally, historically, and politically invested notions of language and language usage, emphasizing that the very existence of '(a) language' is a result of ideological construction and therefore involves power, authority, and control. There is no such thing as a 'neutral' real language; such a neutral notion is one metapragmatic categorization among many, though often the one that indexes power, authority, prestige, and status. Furthermore, one of the essential functions of language is ideological (metapragmatic and indexical) framing: providing contextual cues about who speaks, in what mode, on which topic, and under what circumstances. This ideological function is central to contextualization procedures (Gumperz, 2002), and it may account for much of what Hymes understood as "second linguistic relativity" (1996: 44–45): while linguistic forms may remain stable across contexts, situations, and groups of users, their indexical value may differ quite dramatically. What counts as 'good' language on one occasion may be 'bad' language on another. These insights, as we shall see, have effects on other well-established concepts.

Text

'Text,' oral (face-to-face) as well as written (circulated via inscribed artifacts), is one of the most central notions in the study of language and culture, as the product of language usage and as the central means of cultural transmission. And just like 'language,' it has often been pictured as a reified, referentially transparent, stable object, as something which provides a degree of authenticity in every stage of its existence. Within ethnography, concern for textual performance has a respectable age, and scholars have investigated the poetic patterning of oral texts – 'narrative' or 'folklore' – in attempts to uncover their deeper cultural organization and significance (Hymes, 1996).

Attention to language ideologies has had an impact on at least two aspects of the study of text. The first aspect is the ideological dimension of transmission of texts, and the second is the introduction of indexicality in the analysis of text.

The transmission of Texts As noted above, texts are often seen as stable and capable of unproblematic

transmission. Transmission of texts is often seen as crucial to any understanding of culture – as 'tradition' – and is equally crucial as an ingredient of bureaucratic and institutional practice, in courts, schools, governance, and so forth. In such a transmission, the 'original context' of the text is taken to be carried over to the new occurrence of the text, and impressions of transcontextual stability are created. In discourse analysis, terms such as 'intertextuality' denote such forms of transmission, and they are often used to identify (explicitly) recognizable traces of chunks of discourse across loci of occurrence.

As soon as we adopt a more refined view of context in which indexicality is central, processes of textual transmission become significantly more complex. Let us recall that apart from denotational meanings, texts also carry indexical, metapragmatic values which provide their anchoring into contexts. Every text in context is thus metapragmatically organized in ways that suggest, in context, particular readings or understandings. When texts are transmitted, forms may remain stable – a characteristic called 'entextualization' (Bauman and Briggs, 1990; Silverstein and Urban, 1996b) – but the metapragmatic framing of the text may change. The process by means of which discourses are successively or simultaneously decontextualized and metapragmatically recontextualized depends on such entextualization, so that the forms become a new discourse associated with a new context and accompanied by a particular metadiscourse which provides a sort of 'preferred reading' for the discourse. This new discourse has become a 'text': discourse lifted out of its interactional setting and transmitted together with new suggestion of context (cf. Bauman and Briggs, 1990: 73). Silverstein and Urban (1996b: 1) specified:

> The text idea allows the analyst of culture to extract a portion of ongoing social action – discourse or some nondiscursive but nevertheless semiotic action – from its infinitely rich, exquisitely detailed context, and draw a boundary around it, inquiring into its structure and meaning. This textual fragment of culture can then be re-embedded by asking how it relates to its 'context,' where context is understood as nonreadable surround or background (or if the context is regarded as readable, by asking how the text relates to its 'co-text').

Entextualization is part of what Silverstein and Urban call the 'natural history of discourse.' 'Original' pieces of discourse – socially, culturally, and historically situated unique events – are lifted out of their original context and transmitted into another context, by quoting or echoing them, by writing them down, by inserting them into another discourse, by using them as examples or as data for scientific analysis.

Entextualization builds further on notions of the reflexive nature of language usage (see Lucy, 1993). Through this reflexive dimension, it amends overly linear or static views of context, adding an important praxis-related dimension to text–context relationships. While talking, participants themselves mark those parts of speech that are text and those that are "instructions about how that discourse is to be approached as a text, through replication or with some form of response" (Urban, 1996: 33), e.g., by means of self-corrections, hedges, hesitations, interjections, false starts, explicit qualifications such as 'what I really mean is...' or 'that's not the point.'

Entextualization also connects in relevant ways to another crucial concept: representation. Insofar as representation is always a semiotic act (discursive or nondiscursive, material), and to the degree that representation always involves the 'replication' of an object, a phenomenon, or an event into other modes of existence and other moments of happening, it is a form of entextualization. Hence, representation will also involve metadiscursive qualifications commenting on the ontology or the status of the representation by pointing at distances between the object and the representation, by suggesting that object and representation are identical, or by suggesting that the representation can be treated as identical to the object. We shall see below that this has implications for the history of linguistics.

The Analysis of Texts When indexicality is adopted as one of the primary features and functions of linguistic expression, the analysis of discourse involves far more than an attempt to connect form to content. One of the most important influences from the study of language ideology has been the recognition that formal, poetic patterning in text is indexical, i.e., meaningful in itself. This has an effect on what we understand by 'context' in discourse analysis: the notion of 'contextualization' comes to stand for the generative and dialogical production and interpretation of indexical aspects of language usage (Silverstein, 1992; Gumperz, 2002). Whenever people communicate, they produce forms that fit a particular genre, carry concomitant stylistic features, and thus produce metapragmatic messages about content, direction of interpretation, situatedness in a particular event, social identities, and relationships valid in the event. Utterances are therefore packed with indexical meanings: every utterance is genred, topically organized, linguistically coded, gendered, accented, stylized, and so forth. All of these features occur simultaneously and offer dense clusters of contextualization cues, deployed dialogically in the line-by-line

development of the event (Blommaert, 2005). This is the "dialectical plenitude of indexicality" (Silverstein, 2003: 227; cf. also Goffman's [1974] 'frames within frames').

Thus, in an analysis of marital quarrels and mediation among Zinacantàn Tzotzil speakers in Mexico, John Haviland (1996) showed how the mediator used particular ethnopoetic patterns (parallelisms, refrains, repetitions) in his speech as iconic of the restoration of harmony he was attempting to accomplish. The ethnopoetic patterning here instantiated a genre in which linguistic form iconicized the whole of the event and indexed the roles and relationships produced therein. Similarly, in the analysis of a narrative produced by a refugee from Sierra Leone, Maryns and Blommaert (2001) observed how delicate shifts in (linguistic) code went hand in hand with stylistic and genre shifts, and so indexed affective and epistemic orientations to the places in which particular episodes of the narrative were set. Bundles of formal features, in other words, appeared to organize shifts in orientation towards contextual 'centers' and thus provided new semiotic subject positions, much in the sense of Goffman's changes in footing or Bakhtin's multivocality. And Hanks (1996) showed how a Mayan prayer was indexically structured in such a way as to organize a complex and sequential pattern of participation frameworks. The different structural parts of the prayer oriented towards humans, spirits, space, biographies, and shared histories, and did so in a synthetic, 'polycentric' act of communication.

In all these cases, we see how genred linguistic-narrative form is a crucial contextualizing device and how genre goes hand in hand with identities, roles, topic organization, key, epistemic and affective modes, and so on. Genre, in all of these cases, turned out to be a complex of organized ("ordered," Silverstein, 2003) indexicalities triggering socioculturally presupposable framings for the narratives. Consequently, small, hardly noticeable genre shifts – e.g., from a 'serious' conversation into a 'joke' – indexically reorganize (or reframe) the whole social event, orienting it towards other sets of meanings and expectations of behavior.

There are several implications to this, of which we can only briefly mention a handful here:

1. One implication has to do with the link between communicative-linguistic repertoires and communicative practice. It is clear that both need to be seen in close connection to one another, for certain forms of communicative practice involve the selection of specific linguistic resources from a repertoire (as, for example, the use of a regional accent when quoting someone in a bit of gossip). Thus, the range of stylistic means of speakers may be closely related to their repertoire range, and important forms of inequality may occur in this respect. This is where the study of language ideologies enters the realm of sociolinguistics.
2. Second, there are implications for conceptualizations of form–function relations. Going back to Jakobsonian poetics, we know that contrasts and equivalences, big and small, are crucial indexically. Small formal shifts, sometimes hardly salient linguistically, can have big ideological effects. The phenomenon now often called code-switching can be seen in a new light when we adopt indexicality as a feature of analysis. Often, the shift in (denotational – 'language') code occurs alongside generic and stylistic shifts, and such packages provide complex and layered indexicalities (cf. Rampton, 1995). The actual communicative deployment of code-switching and other forms of contrast and equivalence in linguistic form can be looked at anew.
3. Related to that, discourse analysis can also take a fresh look at poetic patterning in speech, for apart from formal metric organization, talk can also display an indexical rhythm and meter, as Silverstein (2004) demonstrated. Given the 'packaging' of indexicality in formal and structural bundles of features, one may take indexicality as the point of departure rather than as the outcome of formal patterning. This area of investigation awaits focused attention.

Speech community

Apart from the linguistic concepts discussed above, language ideology also dislodges some crucial sociolinguistic concepts. The concept of 'speech community' is one such; its problematic attributions of boundedness, uniformity, and homogeneity have long since been noted. Sociolinguists have long explored the occurrence and distribution of 'languages' (in the artifactual sense outlined above) in societies, and such 'languages' were often accompanied by 'speech communities' in a one-to-one relationship; more sophisticated approaches focused on the occurrence of multiple 'languages' within one 'speech community' or different 'speech communities' sharing one 'language.'

From a language-ideological perspective, two different but related phenomena are at issue. One is the allegiance of groups of people to a denotational code, an artifactual 'language' with a name. The second one is the practical alignment people display towards particular indexical complexes within a stratified system of valuation. The first kind of group would be

called 'language community,' the second one 'speech community' (Silverstein, 1998). There is an obvious relation between both, the first form of groupness being a particular (genred) instance of the second.

To start with the first kind of groupness: this is the kind of allegiance that organizes self-categorizing perceptions and statements of the kind 'I speak X' or 'I am a native speaker of X.' What happens here is an orientation to an artifactual, denotational form of language, usually called by its name (Dutch, English) and imagined as the kind of stable, immanent, clear, and bounded object referred to above. The norm for such a language is codified, and distinctions can be made as to degrees of 'correctness' or knowledge of the ('standard') language thus defined. One 'possesses' this language, and one identifies with an imagined bounded, homogeneous community of 'native speakers' of the language. Thus results an image of singular ethnolinguistic identities – 'I am a speaker of X' – as well as images of static, timeless social formations unified by the language they speak.

The second kind of groupness is more volatile. 'Speech communities' are groups of speakers displaying joint orientations towards complexes of socioculturally presupposable indexicalities which can be used to construct identities and communities. Whereas 'language communities' are ideological constructs, speech communities would be practical constructs, the result of structured social-semiotic practices which, of course, in themselves are based on language ideologies. The existence and actual range of such speech communities depend on the sharedness of indexical values. Consequently, speech communities are shot through with almost every imaginable social variable which can be semiotically indexed. Gender, class, regional background, and spatial trajectories such as migration, ethnicity, levels of education, professions, hobbies, individual biographies, generations, etc. can all be reflected in recognizable semiotic practices and lead to self- or other categorization in particular communities in ways reminiscent of Bourdieu's 'distinction' (1984). Multiple belonging is the rule, and shifts in discourse may signal shifts in orientation towards, or inclusion of, other community-identifying indexicalities.

The language community can be seen as a particular instantiation of 'speech community': a shared orientation towards centered indexicalities, targeting the artifactual language and its explicit norms, often institutionally supported and enforced by the state or other authority-bearing actors. As a point of orientation in discourse, the language community often comes as one layer over other speech community indexicalities. Whenever I teach, for instance, I will talk in Dutch, thus orienting towards a state-organized "regime of language" (Kroskrity, 2000a) which institutionally compels me to use Dutch as the medium of instruction. But the particular register and jargon I use orient towards another community: that of transnational academia in the field of language studies, and my own spoken variety of Dutch will betray my regional, gender, professional, class, and generation belongings as well. Thus we see how through the dense packaging of indexicalities multiple belongings to speech communities crystallize in everyday forms of language usage, making identity issues more a matter of particular combinations of indexical orientations than of categorical identification (Blommaert, 2005: ch. 8).

Silverstein (1998) sketches a series of important implications of these insights. One implication is that debates over 'minority languages' must be seen in a new light, realizing that terms such as 'language loss,' 'language attrition,' or 'language death' may hide a complex reality of shifting orientations of groups towards other indexical centers and thus creating new speech communities (cf. below). The same counts for code-switching, pidginization, and creolization, where the gradual coming into being of a new language community can be seen through the specter of speech communities which serve as the "context of emergence, sustenance and transformation" in the process (Silverstein, 1998: 407). In other words, we stand to gain a far more precise picture of the dynamics of 'enregistering' linguistic codes, varieties, and styles in social processes of identification.

There are other implications as well, and I will mention two important ones. First, taking language ideologies as our point of departure, we can have a new look at language policy. As mentioned above, the artifactual, denotational image of language is the one most often used in institutionalized environments; it is important to realize that this view in itself is the product of language-ideological processes, and that analyses of language policy, consequently, should aim at understanding such processes long before their implementation stage. It is the construction of the ideological image of '(a) language' itself, with the gradual emergence of 'standard' indexical categories for that 'language,' which is the key moment in language policy (cf. the essays in Blommaert, 1999a and Gal and Woolard, 2001a). Analyses based on 'languages' risk accepting the presuppositions of the policy, even when they are criticizing the implementation of the policy. This is particularly important in the context of globalization and the kind of urban multilingualism that results from it. Language policies aimed at 'integrating' new immigrant minorities are most often based on homogeneistic images of language in society, and it is precisely the indexical value of differences

between 'languages belonging here' and 'foreign languages' that provides the dynamic of disqualification and discrimination in education and bureaucracy.

A second, related, implication has to do with ethnolinguistic identity. Identifying people in terms of ethnolinguistic identity – 'I am a speaker of X' – is one of the most common components of lay, bureaucratic, and professional discourses on language in society. Yet, as we have seen, the one-to-one connection between individual speaker, language, and community is untenable. Instead, what counts as 'language' in particular environments, measured and valued against established or emergent regimes of language and in view of categorizablility within the imaginable speech communities, is the issue to be addressed when one intends to understand the relationship between language and identity. This has considerable practical and critical import, because linguistic inequality starts as soon as someone's repertoire is disqualified as 'nonlanguage' or denied the status of 'full languageness,' for instance because of the absence or partial presence of literacy skills or command of the 'standard' or prestige variety of the (artifactually conceptualized) language (cf. Heller, 1999; Jaffe, 1999; also Bourdieu, 1991). Thus, far from ethnolinguistic identity as a 'natural' attribute of humans, what we see is a dynamic of ethnolinguistic categorization, deeply embedded in macrosocial power relations and connected to issues of citizenship, belonging, and enfranchisement.

Language Change

The preceding remarks obviously have a bearing on language change. Silverstein (1979) demonstrated how certain forms of language structure may change due to ideological motivations, i.e., driven by folk understandings of language structure and language use, so that "grammatical change is of a piece with functional-structural [i.e., indexical] change" (Silverstein, 1979: 233). Using examples from English speech acts, Javanese linguistic etiquette, and continental European T/V pronominal systems, he observed how these grammatical patterns developed in the direction of a more strictly denotational system; this, in turn, changed the indexicalities of the grammatical patterns. The move towards a more denotationally transparent system, argued Silverstein, was an effect of the referential ideology of language – people believing that accurate denotation is the primary function of language, and using language accordingly (see also Errington, 1988, 1998 on Javanese).

The ideological drive behind language change was further theorized and elaborated by Irvine and Gal (2000). Focusing on the (mis)recognition of linguistic differences, Irvine and Gal identified three ideological-semiotic effects of processes by means of which language change proceeds: 'iconization,' 'fractal recursivity,' and 'erasure.' Iconization is the process by means of which linguistic features are sensed to be images of the social realities they represent. Thus, language is modeled in such a way that it is seen as iconically representing reality; that is, its indexical signs are understood to correspond to whatever they stand for. Fractal recursivity is the projection of a difference that occurs at one level or in one domain onto another level or into another domain, and erasure is the process by means of which some features of reality are made invisible. Processes of language change driven by language-ideological motivations, they argued, display these processes, either in isolation or in combination with one another.

Irvine and Gal discussed the colonial description of Senegalese languages as an example of the interaction between the three processes. Languages were isolated and projected onto monolingual ethnolinguistic groups and in the process, a geolingual image of language, people, and territory was produced – a case of iconization, graphically represented in ethnolinguistic maps of Senegal. In the process, the multilingualism that characterized many Senegalese communities was made invisible, as it was seen as an effect of earlier historical stages of conquest and colonization – a case of erasure and of fractal recursivity, for this image of multilingualism as an effect of colonialism was a projection of the European colonial impact on Senegal onto precolonial Senegalese society. The general colonial-linguistic enterprise could be seen as a case of iconicization as well: the image of Africans as 'simple' people was mirrored in suggestively 'simple' and straightforward descriptions of the sociolinguistic image of their languages.

A similar case of ideologically driven language change is that of Swahili in the former Belgian Congo. Fabian (1986) described how Swahili was introduced in the Katanga region by the Belgian colonial authorities. Two varieties were introduced: a pidginized variety used in contacts between African workers and Belgian colonial entrepreneurs and administrators, and a 'standard,' sophisticated variety offered only to a small African elite-to-be in advanced education, and tightly controlled by Belgian scholars. Note that, in Irvine and Gal's terms, these two varieties iconicized the segment of the population to which they were offered: the 'simple' pidginized variety was offered to the 'simple' workers, the sophisticated one to the 'sophisticated' elite-to-be. The result on the (socio)linguistic environment was twofold: (1) the introduction of Swahili in a region where it was previously not used, reshuffling the sociolinguistic repertoires of speakers, and (2) the development of

a local variety of Swahili bearing the traces of the pidginized colonial variety, so-called Shaba Swahili. Later, this local variety of Swahili became one of the (semi-)official languages of the independent Congo, and it is now a marker of regional identity and local nationalism – a case of fractal recursivity in which the 'naturalization' of Swahili under colonial rule was adopted by the postcolonial officials and speakers.

What these examples show is the sensitivity of language to wider historical and political-ideological forces. In Silverstein's T/V examples, the gradual rise of macrosocial ideologies of equality was a factor in linguistic development; in the cases discussed by Irvine and Gal and Fabian, colonialism was the wider context within which the construction of new artifactual languages and sociolinguistic images proceeded. They also have a wider critical and metatheoretical effect, because they allow us to revisit established traditions in pragmatics, such as politeness research or speech act theory. What such traditions accomplish is often a *post hoc* and decontextualized theorization of the dominant ideological grids that organize language usage without, however, defining these phenomena as ideological or embedded in communicative practice. Thus, often, either transcendental ('cultural') images of 'norms' are invoked, or ideological and situated processes are presented as simple reflexes of top-down power relations (Agha, 2003). We will return to this below.

The History of Linguistics

The ideological layer in language is a historical layer in which meaning is accomplished by attributing shared and ordered indexicalities to linguistic forms. This counts for folk ideologies of language as well as for specialized professional ones, and the history of the study of language is a history of the formation of ideologies of language as 'discourses of truth' in Foucault's sense: constellations of particular discourses and registers, institutional structures, and professional practices. Foucault himself offered a classic analysis of the emergence of the now dominant view of language as a separate domain of action and knowledge and a neutral, transparent tool for conveying knowledge – the artifactual view we described above (Foucault, 1973). Foucault argued that in the 16th and 17th centuries, a major shift occurred in Western thinking on language, away from a view of language as intrinsically connected to (and embodying) objects and truth. This view was replaced by a new '*épistème*' in which a neutral, instrumental, and autonomous concept of language was seen as the force that enabled the production of modern, rational knowledge.

Bauman and Briggs (2003) further developed and nuanced this thesis. Objecting to Foucault's unified conception of language in the Classical period, they emphasized conflict and hybridity as central to the imagination of language in modernity, thus creating a powerful means of creating social inequality. In Bauman and Briggs's analysis, claims to use of autonomous language become an instrument for claiming modern identities for some by excluding others: the 'folk,' those who speak 'dialects' and perform 'traditions,' the ancestors, the cultural Others. The construction of modern language ideologies has exclusionary, stratifying effects in society. Bauman and Briggs's point of departure is the work of Bacon and Locke, in which elite language is wrested from society and becomes the purified isolated, autonomous object of rationality in modernity, and opposed to 'hybrids' – mixtures of language and society – left to be spoken by the nonelite. Locke adds to this a view of governmentality: "linguistic surveillance becomes a key dimension of [Locke's] pedagogical program" (Bauman and Briggs, 2003: 43), and "tying purification to governmentality rendered language a perfect vehicle for constructing and naturalizing social inequality" (2003: 59).

Bacon and Locke thus provided the language-ideological bedrock of modernity. On that bedrock, several other views were developed, in which the rational and autonomous view of language is contrasted with and used to define modernity's Others: the people of the past, the country folk, people who live in oral cultures and perform 'tradition.' Such views of linguistic Othering were developed by antiquarians such as Aubrey, Bourne, and Brand, whose work tied common people firmly to the past by locating them in a (hybrid) linguistic space of (oral) folklore and tradition. Philologists such as Blackwell, Wood, and Lowth introduced relativism by seeing 'premodern' language as indicative of epochal change. Thus, both contemporary Others and historical Others could be opposed to modern, rational man. Blair's Ossian added a political and universalist-developmental dimension to this set of ideas, showing the path to Herder, who defined poetic (hybrid) form as the essence of national culture. In what at first glance looks as a role reversal, Herder saw poetry (i.e., folklore and tradition) as the 'purest' and most 'natural' expression of the sprit of a *Volk*. In the work of the Brothers Grimm, the rational, autonomous (Lockean) language ideology was used to identify, locate, and appraise folklore. This now more or less finished complex of ideologies was continued, elaborated, and extended in the work of two important early Americanists, Henry Rowe Schoolcraft and Franz Boas, and is thus enshrined in the emergence of American anthropology.

Bauman and Briggs sketched lines of ancestry for modern scientific philological and linguistic concepts

of language, but obviously, the story should not end there. The concrete discursive mechanisms by means of which modern linguistics dissects and produces this view of language through an empirical engagement with linguistic data remain poorly documented (but see e.g., Silverstein, 1996, 2001; essays in Joseph and Taylor, 1990).

A number of studies address linguistics in the historical regime of 'discovery' and colonialism (see Fabian, 1986; Irvine and Gal, 2000 above). This historical era offers rich materials for investigating the connection between the practical work of linguistic scholarship and the larger sociopolitical and cultural-historical schemes in which it could fit (Calvet, 1974; Errington, 2001). In the context of colonial exploration and governance, the colonial subjects' languages were powerful emblems of their inferiority, and their description and categorization consequently displayed all sorts of projections of established Western language ideologies. Rafael's (1993) study of Tagalog under early Spanish rule provided evidence of this dynamics of disqualification, specifically through the projection of Spanish ideologies of literacy onto Tagalog writing, in which we already see articulations of artifactual views of language. Irvine (2001) described how gender and family ideologies were used in the classification of African languages in the 19th century, languages being represented through metaphors of kinship and gender ('daughters,' 'mothers,' etc.), with racial overtones and with hierarchical stratigraphy as a result. Meeuwis (1999) provided an analysis of the way in which Belgian missionaries describing Congolese languages used an 'ideology of the natural' – a view in which languages needed to be established in, kept in, or restored to a 'natural' (God-given) state. This natural state, unsurprisingly, was the one then current in European state nationalisms: the Herderian 'one language–one culture–one territory' complex which defined people as intrinsically monolingual, their language as a '(pure) standard' and their territory as bounded and sociolinguistically homogeneous. These ideological templates provided much of the sociolinguistic imagination of colonial territories, and they were perpetuated in postcolonial language policies and theories of language planning (Blommaert, 1996; Errington, 2001).

The role of linguistics in nationalism has also been rather well documented. The reasons are similar to the ones given for the study of colonialism: the historical era of nation-state building projects offers a fertile basis for investigating the links that exist between particular types of linguistics and larger political-ideological projects. Theorists of nationalism such as Benedict Anderson and Ernest Gellner had already emphasized the importance of print capitalism as an instrument in such projects, but it is clear that more fundamental language-ideological processes were at work as well. Woolard (2004) found traces of early 'modern' imaginings of language as the cement of the nation in 16th-century Spanish treatises, and we have already seen how early Spanish colonizers projected an artifactual view of language onto Tagalog writing (Rafael, 1993). Bauman and Briggs (2003) emphasized the importance of linguistic works such as those of Herder and Grimm in the inception of new European nation-states, but again on a language-ideological basis acquired earlier in the works of Bacon and Locke. A vast area of research awaits attention here.

The role of linguists and other expert actors in 20th-century processes of nation building has drawn substantial attention (see the essays in Blommaert, 1999a; Parlenko and Blackledge, 2002). Bonfiglio (2002) provided a detailed analysis of how Midwestern accents became seen as 'standard' in the early 20th-century United States through the works of important codifiers of American English. Kuzar (2001) discussed the way in which Israeli linguists got inserted in different versions of Zionism, and how their positions in the specter of Zionism correlated their theoretical and methodological preferences; Jaffe (1999) gave detailed attention to the influence of linguists, writers, and translators in the construction of Corsican as a 'minority language.' And Blommaert (1999b) gave an account of the alignment of Tanzanian writers and linguists with the dominant socialist state ideology, leading to a local, strongly politicized form of linguistics and sociolinguistics.

Cultural Variability in Language Ideologies

The study of language ideology grew out of linguistic anthropology and shares the basic preoccupation in this tradition of investigating the nexus of language and culture. It does so by introducing another level of cultural structuring in language: the language-ideological, indexical metalinguistic level. This level drives the development of linguistic structure – the patterns of language change discussed earlier – and it organizes the social, political, and historical framing of language and language use. Consequently, it is a level of significant cultural variability, and some work has started addressing the different organizations of the indexical levels in languages. The outcomes of such work have often included a critique of existing

theory. I shall briefly review some examples of topics and work.

Particular attention has been given to a number of topics. There has been important innovative work on the conceptualization of space among Guugu Yimithirr (Guguyimidjir) and Tzotzil speakers, in which local ideologies of space and of the relationship between gesture and language appear to be influential (Haviland, 1993, 1998). Honorific structures and pronouns have been a topic of inquiry, both as a typical cultural-linguistic category of linguistic occurrences and as forms that display a high sensitivity towards social processes of stratification and categorization. Silverstein (e.g., 1979, 2003) gave frequent examples of pronouns and honorific language use; Errington (1988) was a major study of linguistic etiquette in Javanese; and Irvine (1998) compared honorifics in various African languages. All these authors concluded that linguistic ideologies – perceptions of referential adequacy and conventions of identification – are a factor in explaining the gradual shifts in the systems of honorifics. Such analyses lead to a powerful critique of, for example, politeness theory (as in Errington, 1998). Closely related to this are studies of ritual or ceremonial speech. Here, the performative nature of ritual is at issue, and authors investigate the connections between linguistic form (often formalized, genred speech), occasion, participants, and language ideologies. Hanks (1996) investigated the complex interplay of roles and addressees in Maya shaman prayers; Kroskrity (1998) described how the Kiva speech of the Arizona Tewa involved important language-ideological and identity features; Bauman (1996) inquired into the intricate dynamics of entextualization in Mexican religious ritual festival performances; and Robbins (2001) compared Melanesian and Western Christian ritual discourses in an attempt to arrive at a language-ideological model of ritual. More generally, the emergence and development of genres has also attracted some attention. Kulick (1992) gave detailed evidence of the way in which various genres contribute to the construction of identities in Papua New Guinea; Hill (1998) provided a vivid account of the discourse of nostalgia among Náhuatl speakers; Philips (2000) showed how courtroom discourse in Tonga iconically projects events onto society and the nation-state. Mertz (1998) discussed academic teaching in U.S. law schools as a form of metapragmatic regimentation; Spitulnik (1998a) and Urla (2001) both discussed radio broadcasting, Spitulnik in relation to the complexities of linguistic pluralism in Zambia, Urla in relation to the construction of a counterhegemonic Basque public sphere in Spain. All these authors addressed the way in which specific linguistic forms are deployed in order to attain the ideological effects, and code-switching often emerges as a powerful emblematic feature.

Another topic that has attracted attention is the dynamics of normativity in language. Silverstein's powerful paper (1996) focused on the rise of 'monoglot standard' in the United States; Agha (2003) addressed the development of Received Pronunciation in U.K. English. Errington (1998, 2000) offered perceptive remarks on the tension between the state language and local languages in Indonesia; Blommaert (1999b) described the convergence between state ideology and language practices in Tanzania; Swigart (2000) and Spitulnik (1998b) demonstrated how urban linguistic 'hybrids' such as Urban Wolof (Senegal) and Town Bemba (Zambia) are involved in processes of indexical reorienting towards a more central position in the local speech economies.

All the studies mentioned here contribute to the description of different kinds of relativity – indexical-functional relativity – and provide building blocks for comparative research in this expanding field. Meanwhile, several scholars have addressed phenomena occurring when different contextualization traditions meet, in intercultural communication and translation or in debates over linguistic rights and marginalization. A recent collection of papers in *American Anthropologist* (Mascia-Lees and Lees, 2003) brought the question of language rights into focus and argued convincingly that language rights take on a different shape when looked upon from the point of view of language ideologies (cf. also Silverstein, 1998). Haviland (2003) confronted the linguistic ideologies inscribed in U.S. courtroom practices with the complex process of indexical 'translation' that needs to be undertaken in order to adequately reconstruct the story of Mexican defendants; similar issues of indexical recentering were addressed by Blommaert (2001) in relation to the stories of African asylum seekers in Belgium.

Conceptions of language, coherence, and 'truth' appear to depend heavily on ideologically regimented forms of language use. The forms of regimentation used by, for example, bureaucracy can differ strongly from those used by lay people, even when in established parlance they share a language and live in the same country. This often has important effects of inequality, as Briggs's (1997) analysis of a Venezuelan court case illustrates. Evidently, wherever plurilingual regimes occur, linguistic resources from groups of people can be appropriated and deployed in processes of 'othering.' This happens with languages (Hill, 2001) as well as with accents and speech styles (Rampton, 2003). In the context of globalization processes resulting in heightened perceptual salience

for emblems of identity, such phenomena of indexical recentering within a stratified (unequal) regime may become the key object of theoretical and applied analysis.

Language Ideologies, Norms, and Social Dynamics

The ideological layer of language is a historical layer, as mentioned above, and the historicity of this layer provides for relatively stable trajectories and patterns of indexicality. This fundamental insight allows us to take a different look at 'norms' – an often invoked but rarely theorized linguistic and sociolinguistic notion, often presented as part of 'common knowledge,' 'competence,' or 'intuition,' and generally suggested to be a social convention that comes down on language structure and use.

One of the important contributions of the study of language ideologies is the insight that language always comes with an ideological load which provides comprehensibility through the dynamic of presupposability and inference of indexical meaning, but which also provides a sociopolitical layer of valuation on the utterances produced. Language use, in other words, is intrinsically normative, not in the sense that it always follows established rules, but in the sense that every act of communication will be subject to assessment on grounds of (often implicit) shared complexes of indexicalities – the complexes of indexicalities that provide the basis for speech communities. Norms are patterns of metapragmatic valuation that develop over time in the form of 'enregisterment,' the development of specific forms of language use that carry socially recognizable values and that invite and require continuous interactional re-enactment (Silverstein, 2003).

One way of conceptualizing these patterns of metapragmatic valuation is to start from the nonarbitrary, ordered nature of indexicality. Metapragmatic valuation has a systemic dimension, and by analogy with Foucault's 'orders of discourse' we could speak of 'orders of indexicalities': systemically reproduced indexicalities typically associated with particular genred shapes of language ('enregistered' forms of language use) (Silverstein, 1998, 2003, 2004; Agha, 2003; cf. also Briggs and Bauman, 1992). Whenever people communicate, they display orientations towards such orders of indexicalities, situating their (enregistered) language use in relation to 'norms,' and situating these norms in relation to other norms. Thus, there is always identity work involved, and the orientations towards orders of indexicality are the grassroots displays of 'groupness.' To give an example: young people communicate through orientations to peer group norms; in that way they reproduce the peer group and situate it *vis-à-vis* other peer groups and society at large, thus making the group recognizable both from the inside and from the outside – the particular peer group norms have a specific place in the orders of indexicality to which members orient (Rampton, 1995, 2003).

The systemically reproduced indexicalities are often tied to specific actors, which we can call 'centering institutions' (Silverstein, 1998: 404), often also 'central' actors reproducing the 'doxa' in a particular group. The centering function is **attributive**: it generates indexicalities to which others orient in order to be 'social,' i.e., to produce meanings that 'belong' somewhere and thus to produce categorizable identities. These attributions are emblematic: they revolve around the potential to articulate the perceived 'central values' of a group or system (the 'good' group member, the 'ideal' father/mother/child, 'God,' 'the country/nation,' 'the law,' the 'good' student, the 'ideal' intellectual, the 'real' man/woman...). And this centering almost always involves either perceptions or real processes of homogenization through contrast with other orders of indexicality: orienting towards such a center involves the real or perceived reduction of difference (adopting the enregistered form of language use) and thus the creation of recognizably 'normative' meaning. Centering is thus the semiotic shaping of specific contexts through the creation of contrasts with other contexts. It is the process that generates speech communities.

Centering actors occur at all levels of social life, ranging from the family, through small peer groups, more or less stable communities (e.g. university students, factory workers, members of a church), the state, and transnational communities, all the way to the world system. They are a central feature of what Anderson (1983) called 'imagined communities': though imagined, they trigger specific, enregistered forms of semiotic behaviors and generate groups. But it is worth underscoring that the social environment of almost any individual would by definition be **polycentric**, with a wide range of overlapping and crisscrossing centers to which orientations need to be made, and evidently with multiple 'belongings' for individuals (often understood as 'mixed' or 'hybrid' identities). Furthermore, such environments would be polycentric and **stratified**, in the sense that not every center has equal range, scope, and depth. Small peer groups are not equal to a church community or to the state, and while some centers are normative because of consent (e.g., peer groups), others generate normativity primarily through coercion (e.g., the labor environment or the state in various respects). Consequently, orders of indexicality are

obviously stratified and not all 'loads' have equal value.

Orders of indexicality provide us with a conceptual tool for capturing the moment-by-moment perceptions of stability in language use; they do not, however, suggest total stability or homogeneity because they are subject to permanent re-enactment in situated communicative practices. Agha (2003: 246) argues that, at the lowest level of social practice, norms get transferred through 'speech chains': concrete networks in which people switch communicative roles – from speaker to hearer and vice versa – and so produce a dynamics of communicability and presupposability of norms. Speech chains could be seen as emergent speech networks which display orientations to specific orders of indexicality, and whose communicative practice is the locus of language-ideological innovation. In this way, macrosocial processes can be brought down to the lowest microsociological level of practice.

See also: Anthropological Linguistics: Overview; Linguistic Anthropology.

Bibliography

Agha A (2003). 'The social life of cultural value.' *Language & Communication 23,* 231–273.

Anderson B (1983). *Imagined communities.* London: Verso.

Bauman R (1996). 'Transformations of the word in the production of Mexican festival drama.' In Silverstein & Urban (eds.). 301–327.

Bauman R & Briggs C (1990). 'Poetics and performance as critical perspectives on language and social life.' *Annual Review of Anthropology 19,* 59–88.

Bauman R & Briggs C (2003). *Voices of modernity: language ideologies and the politics of inequality.* Cambridge: Cambridge University Press.

Blommaert J (1996). 'Language planning as a discourse on language and society: the linguistic ideology of a scholarly tradition.' *Language Problems & Language Planning 20,* 199–222.

Blommaert J (ed.) (1999a). *Language ideological debates.* Berlin: Mouton de Gruyter.

Blommaert J (1999b). *State ideology and language in Tanzania.* Cologne: Rüdiger Köppe.

Blommaert J (1999c). 'The debate is open.' In Blommaert (ed.). 1–38.

Blommaert J (2001). 'Investigating narrative inequality: African asylum seekers' stories in Belgium.' *Discourse and Society 12,* 413–449.

Blommaert J (2005). *Discourse: a critical introduction.* Cambridge: Cambridge University Press.

Bonfiglio T (2002). *Race and the rise of Standard American.* Berlin: Mouton de Gruyter.

Bourdieu P (1984). *Distinction: a social critique of the judgement of taste.* London: Routledge.

Bourdieu P (1991). *Language and symbolic power.* Cambridge: Polity Press.

Briggs C (1986). *Learning how to ask.* Cambridge: Cambridge University Press.

Briggs C (1997). 'Notes on a "confession": on the construction of gender, sexuality, and violence in an infanticide case.' *Pragmatics 7,* 519–546.

Briggs C & Bauman R (1992). 'Genre, intertextuality and social power.' *Journal of Linguistic Anthropology 2,* 131–172.

Calvet L-J (1974). *Linguistique et colonialisme: petit traité de glottophagie.* Paris: Payot.

Cameron D (1995). *Verbal hygiene.* London: Routledge.

Errington J (1988). *Structure and style in Javanese: a semiotic view of linguistic etiquette.* Philadelphia: University of Pennsylvania Press.

Errington J (1998). *Shifting languages: interaction and identity in Javanese Indonesia.* Cambridge: Cambridge University Press.

Errington J (2000). 'Indonesian's authority.' In Kroskrity (ed.). 205–227.

Errington J (2001). 'Colonial linguistics.' *Annual Review of Anthropology 30,* 19–39.

Fabian J (1986). *Language and colonial power.* Cambridge: Cambridge University Press.

Foucault M (1973). *The order of things.* New York: Vintage.

Gal S & Woolard K (2001a). 'Constructing languages and publics: authority and representation.' In Gal & Woolard (eds.). 1–12.

Gal S & Woolard K (eds.) (2001b). *Languages and publics: the making of authority.* Manchester: St. Jerome.

Goffman E (1974). *Frame analysis: an essay on the organization of experience.* New York: Harper & Row.

Gumperz J (2002). 'Response essay.' In Eerdmans S, Prevignano C & Thibault P (eds.) *Language and interaction: discussions with John J. Gumperz.* Amsterdam: John Benjamins. 105–126.

Hanks W (1996). 'Exorcism and the description of participant roles.' In Silverstein & Urban (eds.). 160–200.

Haviland J (1993). 'Anchoring, iconicity and orientation in Guugu Yimithirr pointing gestures.' *Journal of Linguistic Anthropology 3,* 3–45.

Haviland J (1996). '"We want to borrow your mouth": Tzotzil marital squabbles.' In Briggs C (ed.) *Disorderly Discourse.* New York: Oxford University Press. 158–203.

Haviland J (1998). 'Early pointing gestures in Zincantan.' *Journal of Linguistic Anthropology 8,* 162–196.

Haviland J (2003). 'Ideologies of language: some reflections on language and US law.' *American Anthropologist 105,* 764–774.

Heller M (1999). *Linguistic minorities and modernity: a sociolinguistic ethnography.* London: Longman.

Hill J (1998). '"Today there is no respect": Nostalgia, "respect," and oppositional discourse in Mexicano (Nahuatl) language ideology.' In Schieffelin *et al.* (eds.). 68–86.

Hill J (2001). 'Mock Spanish, covert racism and the (leaky) boundary between public and private spheres.' In Gal & Woolard (eds.). 83–102.

Hymes D (1996). *Ethnography, linguistics, narrative inequality: toward an understanding of voice.* London: Taylor & Francis.

Irvine J (1998). 'Ideologies of honorific language.' In Schieffelin *et al.* (eds.). 51–67.

Irvine J (2001). 'The family romance in colonial linguistics: gender and family in nineteenth-century representations of African languages.' In Gal & Woolard (eds.). 13–29.

Irvine J & Gal S (2000). 'Language ideology and linguistic differentiation.' In Kroskrity (ed.). 35–83.

Jaffe A (1999). *Ideologies in action: language politics on Corsica.* Berlin: Mouton de Gruyter.

Joseph J & Taylor T (eds.) (1990). *Ideologies of language.* London: Routledge.

Keane W (2003). 'Semiotics and the social analysis of material things.' *Language & Communication 23,* 409–425.

Kroskrity P (1998). 'Arizona Tewa Kiva speech as a manifestation of a dominant language ideology.' In Schieffelin *et al.* (eds.). 103–122.

Kroskrity P (2000a). 'Regimenting languages: language ideological perspectives.' In Kroskrity (ed.). 1–34.

Kroskrity P (ed.) (2000b). *Regimes of language.* Santa Fe: SAR Press.

Kulick D (1992). *Language shift and cultural reproduction.* Cambridge: Cambridge University Press.

Kuzar R (2001). *Hebrew and Zionism: a discourse analytic cultural study.* Berlin: Mouton de Gruyter.

Lucy J (ed.) (1993). *Reflexive language: reported speech and metapragmatics.* Cambridge: Cambridge University Press.

Maryns K & Blommaert J (2001). 'Stylistic and thematic shifting as a narrative resource: assessing asylum seekers' repertoires.' *Multilingua 20,* 61–84.

Mascia-Lees F & Lees S (eds.) (2003). 'Language ideologies, rights and choices: dilemmas and paradoxes of loss, retention, and revitalization.' *American Anthropologist 105(4),* 710–781.

Meeuwis M (1999). 'Flemish nationalism in the Belgian Congo versus Zairian anti-imperialism: continuity and discontinuity in language ideological debates.' In Blommaert (ed.). 381–423.

Mertz E (1998). 'Linguistic ideology and praxis in US law school classrooms.' In Schieffelin *et al.* (eds.). 149–162.

Milroy J & Milroy L (1985). *Authority in language: investigating language prescription and standardisation.* London: Routledge and Kegan Paul.

Parlenko, Aneta & Adrian Blackledge (eds.) (2002). 'Language ideologies in multilingual contexts.' *Multilingua 21(2–3),* 121–302.

Philips S (2000). 'Constructing a Tongan nation-state through language ideology in the courtroom.' In Kroskrity (ed.). 229–257.

Rafael V (1993). *Contracting Colonialism: Translation and Christian conversion in Tagalog society under early Spanish rule.* Durham: Duke University Press.

Rampton B (1995). *Crossing: language and ethnicity among adolescents.* London: Longman.

Rampton B (2003). 'Hegemony, social class and stylisation.' *Pragmatics 13,* 49–83.

Robbins J (2001). 'Ritual communication and linguistic ideology.' *Current Anthropology 42,* 591–641.

Schieffelin B, Woolard K & Kroskrity P (eds.) (1998). *Language ideologies: practice and theory.* New York: Oxford University Press.

Silverstein M (1976). 'Shifters, linguistic categories and cultural description.' In Basso K & Shelby H (eds.) *The meaning of meaning.* Albuquerque: University of New Mexico Press. 11–55.

Silverstein M (1979). 'Language structure and linguistic ideology.' In Clyne P, Hanks W & Hofbauer C (eds.) *The elements: a parasession on linguistic units and levels.* Chicago: Chicago Linguistic Society. 193–247.

Silverstein M (1992). 'The indeterminacy of contextualization: when is enough enough?' In Auer P & Di Luzio A (eds.) *The contextualization of language.* Amsterdam: John Benjamins. 55–76.

Silverstein M (1996). 'The secret life of texts.' In Silverstein & Urban (eds.). 81–105.

Silverstein M (1998). 'Contemporary transformations of local linguistic communities.' *Annual Review of Anthropology 27,* 401–426.

Silverstein M (2001). 'From the meaning of meaning to the empires of the mind: Ogden's orthological English.' In Gal & Woolard (eds.). 69–82.

Silverstein M (2003). 'Indexical order and the dialectics of social life.' *Language & Communication 23(3–4),* 193–229.

Silverstein M (2004). 'Cultural' concepts and the language-culture nexus. *Current Anthropology 45,* 621–652.

Silverstein M & Urban G (eds.) (1996a). *Natural histories of discourse.* Chicago: University of Chicago Press.

Silverstein M & Urban G (1996b). 'The natural history of discourse.' In Silverstein & Urban (eds.). 1–17.

Spitulnik D (1998a). 'Mediating unity and diversity: the production of language ideologies in Zambian broadcasting.' In Schieffelin *et al.* (eds.). 163–188.

Spitulnik D (1998b). 'The language of the city: Town Bemba as urban hybridity.' *Journal of Linguistic Anthropology 8,* 30–59.

Swigart L (2000). 'The limits of legitimacy: language ideology and shift in contemporary Senegal.' *Journal of Linguistic Anthropology 10,* 90–130.

Urban G (1996). 'Entextualization, replication, and power.' In Silverstein & Urban (eds.). 21–44.

Urla J (2001). 'Outlaw language: creating alternative public spheres in Basque free radio.' In Gal & Woolard (eds.). 141–163.

Woolard K (1998). 'Introduction: language ideology as a field of inquiry.' In Schieffelin *et al.* (eds.). 3–47.

Woolard K (2004). 'Is the past a foreign country? Time, language origins, and the nation in early modern Spain.' *Journal of Linguistic Anthropology 14,* 57–80.

Language in Computer-Mediated Communication

N Baym, University of Kansas, Lawrence, KS, USA

Before computers were networked, the terms 'language' and 'computer' only occurred together in describing programming code. In recent years, however, the Internet and related technologies have become popular and pervasive media for human communication. Since the late 1970s (e.g., Hiltz and Turoff, 1978), questions of how human language is used in these media have become important concerns for linguists, communication scholars, sociologists, and researchers from other disciplines. This article begins by describing the primary forms of computer-mediated communication (CMC), as well as the characteristics that differentiate these media from one another. It then addresses the primary issues that have dominated the discussions of computer-mediated language. A brief caveat is in order. The Internet changes so rapidly that the forms of CMC available, and the ways in they are used and by whom, evolve more quickly than print forms such as this article can accommodate. By the time this article reaches its readers, new forms may well have emerged.

Characteristics of Computer Language Media

In the simplest sense, CMC can be distinguished on the basis of whether it occurs in real-time (synchronous CMC) or whether there is a time delay between messages (asynchronous CMC). With synchronous forms, participants type messages that appear on others' screens as soon as they are sent. Readers reply instantly, leading to a rapid exchange of brief messages. Chat rooms, where multiple people gather on the basis of shared interests, are one form of synchronous CMC, as is instant messaging, which usually occurs between two people. Synchronous CMC is also found in richly visual graphical online games (including multi-player role-playing games in which people play together to achieve goals such as gaining treasure or slaying mythical monsters) and as an accompaniment to online versions of traditional games such as poker, dominoes, and backgammon. Multi-user domains, or MUDS, are textual online 'spaces' designed for functions as varied as role-playing, generalized socializing, and education. Like other synchronous forms of CMC, users of a MUD interact in real time if they are online simultaneously and are in the same 'room' or area of virtual space.

The most popular forms of CMC are asynchronous. Email, the electronic equivalent of letters, may be person-to-person, or may be group communication organized through mailing lists based on interest, family, friendship, work, or other ties. Websites are also asynchronous, with static information posted for readers. On weblogs (or 'blogs'), a writer or collective of writers posts comments on a daily or even hourly basis. Readers can often engage in discussion of those comments through hyperlinked sections that appear as second pages. Webboards, bulletin boards, and newsgroups are asynchronous forms of topical discussion in which large groups of users may gather on their own time to read and leave messages. Though the distinction based on synchronicity is common, in practice the lines between the two may blur. If people are online simultaneously and checking often, email can function in a near synchronous manner. Short message services (SMS), also known as text messaging, send very brief messages through mobile phones in exchanges that can happen quickly enough to be considered synchronous or with enough lag between turns to render them asynchronous.

Computer-mediated language can also be characterized by a set of other features that can influence interaction. These characteristics have also proven helpful for thinking about how forms of CMC differ from one another and how they compare to other forms of language use such as face-to-face communication (the usual standard for comparison in CMC research), telephone calls, and traditional letters (see the discussion of oral and written language below). In contrast to most other media forms, CMC can be either one-to-one or one-to-many. Email, instant messaging, and SMS are usually directed at single users, while mailing lists, chat rooms, and discussion boards are usually written for a broad and often semi-anonymous audience. The ability for each Internet user to communicate one-to-many is an unprecedented transformation in the landscape of communication media, which have historically limited this capacity to a small number of influential mass media producers.

Forms of computer-mediated language media also vary in the degree of interactivity, or ability to respond to messages, that they offer. Email and instant messaging are highly interactive, as messages are generally meant to invoke responses (one-to-many email as occurs on mailing lists and with spam is an exception). Websites, on the other hand, may not offer any means for readers to respond to their contents. A single form, such as the weblog, may vary in the degree of interactivity it offers depending on whether the author includes the means for readers to leave comments or reply to those comments that are left.

In strong contrast to face-to-face interaction and to a lesser extent the telephone, computer-mediated language media seem to transcend space. There are many locations where Internet access is difficult or impossible, which affects which languages are common online (more on this below). However, so long as people are online, messages exchanged across continents are indistinguishable from those sent between rooms in the same building, both in form and speed of transmission. This can create a sense of being close to one another. Since sending messages via computer is generally less expensive than using long distance telephone (let alone travel), computer mediation can greatly increase the opportunities people have to use language with those far away.

By virtue of being electronic, computer-mediated language can be stored and replicated. Although forms of CMC vary in how ephemeral they may feel, most, if not all, are replicated as they are transmitted and stored on servers and backups. Most forms of CMC can be copied and sent on to others. Even messages sent person-to-person can be retrieved and replicated for other individuals or large audiences. This was the case when all email sent at the failed corporation Enron was posted to the World Wide Web as part of court proceedings. Chat and instant messaging seem to pass as soon as they occur, yet they can be logged. However, despite the inherent replicability and storage potential of computer-mediated language, storing and archiving examples of online language use has proven to be a difficult logistical challenge that has become a source of particular concern for librarians and information scientists. Computer-mediated language sites that are present one day may be gone or transformed the next.

Some forms of online language are more hypertextual than others. In hypertextual CMC, exemplified by websites, particular words or phrases on one page are linked to other pages with more information. Clicking on the hyperlink allows users to follow up on those bits that interest them most. In contrast to more linear forms of CMC such as chat, gaming, and instant messaging, in hypertextual CMC, information is connected through multiple possible relationships, giving readers the choice of what to pursue and allowing different readers of the same text potentially very different reading experiences. Email, which was once linear, has become more hypertextual as the ability to embed clickable links within messages has become common.

Finally, CMC media differ in richness (Daft and Lengel, 1984), or the extent to which nonverbal cues are apparent. In its early days, all forms of CMC were extremely lean, as all messages were restricted to the same monochrome font in the limited ASCII character set. The ability to format messages by varying fonts and colors and embedding images and sounds has made many forms of CMC, especially websites and graphic interfaces, richer. When CMC is lean, as it often is in more text-based forms, the reduction of physical appearance cues, along with the evidence of status and attractiveness they bear, can create a kind of invisibility or anonymity. Leanness also positions language as the sole means of creating social context among interactants.

With this background, the article now turns to key issues that have dominated the research on computer-mediated language. These include the primacy of English versus other languages, comparisons between CMC and speech and writing, online playfulness, hostility, and the influence of gender on online language use. The essay concludes with a consideration of the influence online language use is likely to have on offline language.

Languages Used in Computer-Mediated Communication

Measuring the use of any particular language in online communication is challenging given the variety of forms of CMC and their varying accessibility. Efforts to ascertain language use have usually estimated the proportion of Internet users who speak a given language. English emerges as the most-used language on computer networks, although the proportion of English speakers relative to those who speak other languages on the Internet has fallen considerably since the late 1990s. The business Translate to Success has compiled data from varied surveys of Internet users to estimate that in 2004, 38.3% of Internet users spoke English. Chinese, Japanese, and Korean are also popular, constituting 11.2%, 10%, and 4.1%, respectively. The most common European languages are German (6.8%), Spanish (5.5%), Italian (3.9%), French (3.5%), Portuguese (3.1%), Russian (3%), and Dutch (2%). Only 1% of the world's Internet users speak Arabic. Fewer than 0.1% speak any African language. An effort to conduct a language census of weblogs (blogcensus.net) has resulted in the indexing of over two million blogs. More than half of these are in English, followed in dramatically smaller numbers by those in French, Portuguese, and Farsi (each of which occurred in between 10 000 and 90 000 blogs). Polish, German, Spanish, and Italian are the only other languages found in more than 10 000 blogs each.

That language use on the Internet maps so poorly onto language use in the global population reflects the wealth required to incorporate computer-technology into a populace, as well as historical and governmental constraints on the implementation of the

technology in given cultures. Furthermore, in some nations, such as India, English may be the language used to access the Internet, even if it is not the primary language spoken by the majority of its citizens. The overrepresentation of languages used in wealthy countries, especially English, has often given rise to a sentiment that the Internet represents a further colonization of poor nations by those with greater wealth, particularly the United States.

The use of a given language on the Internet is also affected by the technology itself. Until recently, online writing was restricted to the ASCII character set, which is designed exclusively for the Latin alphabet. With the advent of Unicode, people can now write with other alphabets; however, this technology is neither available to nor used by all. The result has sometimes been considered a form of "typographical imperialism" (Herring and Danet, 2004) with potential social, political, economic, and linguistic consequences. For instance, Greece has seen recent controversy surrounding what is called Greeklish, an online version of written Greek in which the Latin alphabet is used to transliterate (Koutsogiannis and Mitsihopoulou, 2004). While some Greeks, particularly the young, view this as a means of participating in the global discourse of the Internet, others view it as a threat to the integrity and history of the Greek language.

Writing, Speaking, and Hybridity

Scholars striving to describe the linguistics of CMC have often done so by comparing online interaction to face-to-face communication and writing. Although in earlier work, the comparison presupposed an erroneous assumption that all computer-mediated language was the same, other scholars have limited their comparisons to particular forms of CMC. Biber (1986), Chafe (1982), and Mulkay (1985, 1986) are among those who have compared oral and written language. In general, scholars examining diverse forms of CMC in a number of languages (particularly English, French, Swedish, and Norwegian) have found that CMC resembles both written language and oral conversation (Baron, 2000; Baron and Ling, 2003; Baym, 1996; Danet, 1997; Ferrera *et al.*, 1991; Hård af Segerstad, 2005; Herring, 2001; Ling, 2005). CMC is like writing in many ways. The text usually bears an address. Messages can be edited prior to transmission. The author and reader are usually geographically (and often temporally) separated, messages can often be read by anonymous readers who may not respond. It is not possible for interlocutors to overlap one another nor to interrupt. Context must be created through the prose so that messages are often explicit and complete. There is rarely an assumption of shared physical context. Messages are replicable and can be stored. Vocabulary, syntax, spelling, and the use of uncontracted forms may make online interaction more like writing than speech.

On the other hand, there are many ways in which online language better resembles speech. Messages are generally related to prior ones, often through turn-taking, although disrupted turn adjacency and lack of feedback can render turn-taking challenging (Herring, 2001). Messages are based in a relationship between writer and reader. There is often a history of shared referents and speech conventions (e.g., Hymes, 1986). The audience is usually able to respond and often does so quickly, resulting in reformulations of original messages. Topics change rapidly. The discourse often feels ephemeral, and often is not stored by recipients despite the capacity for storage. Furthermore, online language can be marked by colloquial and nonstandard spellings that foreground phonetic qualities (e.g., 'gotta' instead of 'have to').

Ultimately, many scholars conclude that online language is an "interactive written register" (Ferrera *et al.*, 1991), hybrid (Danet, 1997), creole (Baron, 1998), or "uncooked linguistic stew" (Baron and Ling, 2003) that blends elements of written and oral language with features that are distinctive to this medium, or at least more common online than in any other language medium. Among the most commented upon features of online writing are the use of abbreviation (e.g., TTYL for 'talk to you later'), the use of asterisks as brackets to simulate underlines, and upper case lettering to indicate emphasis. A number of deletions have also been noted, including the deletion of subject pronouns (e.g., 'gotta go now'), vowels, and punctuation. In SMS, spaces may be deleted as well, and there may not be any adjectives or adverbs (Hård af Segerstad, 2005; Ling, 2005). Messages in instant messaging, chat, and SMS are considerably shorter than those in conversation, writing, or most other forms of CMC, and in some synchronous forms of CMC such as IM, utterances may be broken up across turns.

In addition to the variety of forms of CMC, the task of identifying distinguishing characteristics of CMC is further complicated by the fact that there are so many other factors that influence language use in any medium (Baron and Ling, 2003; Baym, 1995a; Herring, 2001). As Basso (1974) has noted, written language always depends on its cultural and social contexts, functions, and the relationship between writer and reader. Each of these factors is at play in determining the linguistic form of any single instance

of computer-mediated language use. Furthermore, individuals have distinct personal styles. As seen below, gender also influences language style.

In group CMC contexts where many people hold discussion over long periods of time, group norms develop that influence language use. These may include distinctive words, such as 'whuggle,' a form of greeting documented by Cherny (1999) or 're,' a short form of 'hello again' used to greet for a second time in an Internet Relay Chat channel (Werry, 1996). In some cases, groups develop entire genres of language use distinctive or tailored to their forum, such as humorous lists of 'unasked questions' in a soap opera discussion (Baym, 1995b). Stylistic standards of formality and politeness also emerge over time within particular groups.

Playfulness in Computer-Mediated Language

Some of the distinctive linguistic features found in CMC can be attributed to a desire to save keystrokes, especially features such as deletions in synchronous interaction. However, many of the phenomena that occur in online language are better described as playful than time-saving. A number of scholars, most notably Danet (2001), have noted the playfulness of CMC. Danet (2001) argues that the play found on CMC, especially in synchronous CMC, can be traced to several influences. Interactivity and instant feedback can lead to a sense of immersion that may facilitate play. The lack of clear authorities and formal governing structure of online spaces can promote playfulness. The legacy of hacker culture with its love of wordplay, puns, irony, flippancy, and play with typography and spelling is also an influence as hackers were early adaptors of the Internet. Finally, the anonymity of many online spaces can be conducive to play since the enhanced role of language in creating an identity offers participants greater latitude than normal in presenting a self. For instance, many forms of online interaction require as a starting point that interactants choose a nickname or screen name by which they will be known.

Other reasons for the playfulness of CMC may be a desire for spontaneity or to increase the socioemotional appeal of messages in this nonverbally impoverished medium. Many of the 'errors' that make online interaction more like speech than writing can be understood as efforts to create a friendly, informal, conversational tone, which in turn gives rise to further playfulness. Following Bauman (1975), Baym (1995b) argues that online language often serves as a mode of performance that strives to entertain, to enhance the appeal of the discourse, to build identities for the performers, and to foster friendly relationships among interactants.

Among the most common forms of linguistic play in CMC are the aforementioned acronyms. Often-used acronyms including BTW (by the way), IMHO (in my humble opinion), and many others are compiled into online dictionaries posted on the Web. In particular, CMC contexts, people often develop unique playful acronyms such as PVGD, used in a soap opera discussion group to explain why the children in the fictional town of Pine Valley aged so rapidly (the GD stands for 'glandular disorder'), or YMMV for 'your mileage may vary' offered when sharing personal experience in an online discussion group for pregnant women. Creative acronyms may be used to substitute for nonverbal expressions, as in LOL for 'laughing out loud.' Language play can also be seen in the development of new words in online contexts. Spam (unwanted barrages of messages), flaming (hostility), and blogging (writing online diaries) are examples of such words that have made it out of the Internet into wider discourse. Many have noted how common humor is in online contexts (e.g., Baym, 1995b; Myers, 1987).

Online Hostility

In contrast to theories that CMC fosters a playful affiliative environment, there has long been speculation that because of its potential anonymity, or reduction of nonverbal cues, CMC would foster the use of more hostile language, or flaming. Early research (e.g., Kiesler *et al.*, 1984) argued that because the computer medium is lean, it creates an environment in which norms are suspended. As a consequence, there would be more swearing, insults, name calling, and expression of negative affect. This (error-filled) comment from a group discussion of the television show *Star Trek*, is a typical flame: "Will you stupid jerks get a real life. Everyone with half a brain or more know that a human and a Kligon can not mate. The Klingon mating procedure would kill any human (except one with a brain like you). Stay off the net stoopid!" In group interactions, such messages usually lead to what has become known as a flame war. Such flame wars follow a fairly predictable pattern in which the original flame is followed by an equally hostile retort. The hostilities escalate, drawing in more participants. Other participants then chime in urging the original participants to move the discussion off-list or to ignore the hostilities. Eventually people lose interest and the discussion dies out. Many sources on the Internet can be found describing this pattern and offering 'netiquette' tips to

prevent flame wars. Flaming is also concern in person-to-person interactions online. Fearing flaming, some organizations now require that some forms of communication, such as evaluations and reprimands, be conducted face-to-face rather than online in order to prevent inappropriate hostility.

One explanation of flaming is the reduced "social presence" (Short *et al.*, 1976) of other interactants given the leanness of the computer medium. Other scholars (Lea *et al.*, 1992) have argued that flaming cannot be explained by the medium alone, since the amount of hostility varies tremendously across online contexts. They argue instead for a normative explanation, in which some groups create contexts that encourage more argumentative styles. Flaming has also been linked to masculinity, or "the chest-thumping display of online egos" (Myers, 1987: 241).

Despite the common sense that flaming is rampant in CMC, research suggests that it is perceived as more common than it actually is. Rice and Love (1987) found in their study of an organization's computer-mediated messages that although socioemotional expression was common, only 0.2% of the messages were antagonistic. Lea *et al.* (1992) theorize that flaming may seem more common than it is because single messages may be seen by so many people and because hostile messages are so memorable. In fact, people are considerably more likely to be nice than to flame in online contexts (e.g., Preece and Ghozati, 1998).

Gender

The speculation that flaming is associated with gender is indicative of a strong focus on gender that has characterized much of the linguistic work on CMC. Early CMC rhetoric often suggested that the elimination of the nonverbal cues to gender would result in the elimination of gender from language use, with optimistic consequences ranging from women's voices being granted more status to the elimination of gender as a meaningful cultural category. In some online contexts, gender becomes a subject for linguistic play. One much-studied MUD, Lambda MOO, offers participants nine gender options for their identity, each with its own set of pronouns (Danet, 2001). In addition to male, female, and neuter, people can choose to identify as: either, Spivak, splat, plural, egotistical, royal, or 2nd person. Third person descriptions of each of these options would be he, she, it, s/he, E, e*, they, I, we, you, and per, respectively.

However, research has consistently investigated whether the differences often noted in face-to-face conversation persist in CMC and whether gender can be identified from language style. Empirical findings are mixed, but do suggest modest gender differences. In discourse analytic studies of asynchronous discussions, Herring (2001) argues that men use an adversarial style, and women a supportive/attenuated style oriented toward affiliation. Although there is more overlap than difference in the messages written by men and women, Herring asserts that at the extremes, longer messages, strong assertions of opinions, and crude language are more likely to be written by men. One MUD described by Kendall (2002) describes the aggressive style of answering questions seen often within the group as MAS for 'Male Answer Syndrome.' Men are also more likely to open and close conversations. Sexist language has also been found in CMC (e.g., Gurak, 1997). Kendall (2002), for instance, describes how women outside the group are depicted as sexual objects. When someone mentions seeing a woman, for instance, a typical response is "did you spike 'er?" (Kendall, 2002: 85).

Messages written by women are more likely to be short and to include qualifications, justifications, apologies, and expressions of support (Herring, 2001). Evidence for a female affiliative style can also be seen in the finding that women's IM closings take twice as many turns and are nearly three times as long as male closings. Women are also nearly three times more likely to begin SMS interactions with openings (Baron and Ling, 2003).

However, while Savicki *et al.* (1997) found that groups in which there were more men used more factually oriented language and calls for action, less self-disclosure, and fewer attempts at tension prevention and reduction, they did not find any gender differences in expressions of opinion, apology, questions, or even flaming. Ling (2005) finds that women's SMS messages are longer, exhibit more complex structure, and are more like writing than men's. In their statistical analysis, Savicki *et al.* (1997) report that gender had only modest explanatory power in differentiating the language style of messages.

Gender is a factor not just in message style, but also in perception of message style, with impacts on response (or nonresponse) to those messages. Herring (1996) argues that men and women differ in their perception of aggressive messages. While men are more likely to see them as evidence of freedom of speech, candor, and healthy debate, women are more likely to see them as hostile and unconstructive.

The Effect of Computer-Mediated Language on Offline Language Use

CMC offers a hybrid, playful, and occasionally confrontational style of language that closely resembles language in other media, yet has some distinguishing

characteristics. As is often the case when new language forms emerge, some develop concern that standard language is under threat. Newspaper articles have worried, for instance, that the brief exchanges of instant messaging will lead to an inability to conduct face-to-face conversations or that nonstandard spelling and punctuation will decimate grammar as we know it. Although there have been few if any studies directly addressing this phenomenon, the evidence from studies of computer-mediated language do not offer strong reasons for concern. CMC has been shown to have depended on a variety of influences, including the technology, the purpose of the interaction, the norms of the group, the communication style of the speakers' social groups offline, and the idiosyncrasies of individuals. There is no consistent standardized 'Computer Mediated Language.'

Baron and Ling's (2003) analysis suggests there are actually far fewer such deviations from standard language forms than people would expect given the attention they are paid. Furthermore, only a few of the nonstandard features of language that do occur in CMC are due to inattention or lack of awareness of standards (Herring, 2001). Most are instead deliberate choices made as language users adapt to the technical and social contexts of their interactions. The evidence thus suggests that people are aware that different situations call for different language styles, and are unlikely to lose the ability to hold face-to-face conversations, spell, or use punctuation as a result of time spent online.

See also: E-mail, Internet, Chatroom Talk: Pragmatics; Computer-Mediated Communication: Cognitive Science Approach.

Bibliography

Baron N S (1998). 'Letters by phone or speech by other means: the linguistics of email.' *Language and Communication 18,* 133–170.

Baron N S (2000). *Alphabet to email: how written English evolved and where it's heading*. London: Routledge.

Baron N & Ling R (2003). 'IM and SMS: a linguistic comparison.' Paper presented at the Association of Internet Researchers, Internet Research 4.0. Toronto, October.

Basso K (1974). 'The ethnography of writing.' In Bauman R & Sherzer J (eds.) *Explorations in the ethnography of speaking*. Cambridge: Cambridge University Press. 425–432.

Bauman R (1975). 'Verbal art as performance.' *American Anthropologist 77,* 290–311.

Baym N (1995a). 'The emergence of community in computer-mediated communication.' In Jones S (ed.) *Cybersociety: computer-mediated communication and community.* Thousand Oaks, CA: Sage. 138–163.

Baym N (1995b). 'The performance of humor in computer-mediated communication.' *Journal of Computer-Mediated Communication* 1(2). http://www.ascusc.org/jcmc/vol1/issue2/baym.html.

Baym N (1996). 'Agreement and disagreement in a computer-mediated group.' *Research on Language and Social Interaction 29,* 315–346.

Biber D (1986). 'Spoken and written textual dimensions in English: resolving the contradictory findings.' *Language 62,* 384–414.

Chafe W (1982). 'Integration and involvement in speaking, writing and oral literature.' In Tannen D (ed.) *Spoken and written language: exploring orality and literacy.* Norwood, NJ: Ablex. 35–54.

Cherny L (1999). *Conversation and community: chat in a virtual world.* Stanford: CSLI Publications.

Daft R L & Lengel R H (1984). 'Information richness: a new approach to managerial behavior and organizational design.' *Research in Organizational Behavior 6,* 191–233.

Danet B (1997). 'Books, letters, documents: the changing aesthetics of texts in late print culture.' *Journal of Material Culture 2(1),* 5–38.

Danet B (2001). *Cyberpl@y: communicating online.* Oxford: Berg.

Ferrara K, Brunner H & Whittemore G (1991). 'Interactive written discourse as an emergent register.' *Written Communication 8,* 8–34.

Gurak L L (1997). *Persuasion and privacy in cyberspace: the online protests over Lotus MarketPlace and the clipper chip*. New Haven: Yale University Press.

Hård af Segerstad Y (2005). 'Language use in Swedish mobile text messaging.' In Ling R & Pederson P (eds.) *Mobile communications: renegotiation of the social sphere*. London: Springer. 313–333.

Herring S (1996). 'Posting in a different voice: gender and ethics in computer-mediated communication.' In Ess C (ed.) *Philosophical approaches to computer-mediated communication.* Albany: SUNY Press. 115–145.

Herring S (2001). 'Computer-mediated discourse.' In Schiffrin D, Tannen D & Hamiton H E (eds.) *The handbook of discourse analysis*. Malden, MA: Blackwell. 612–634.

Herring S & Danet B (2003). 'The multilingual Internet.' *Journal of Computer-Mediated Communication 9(1).*

Hiltz S R & Turoff M (1978). *The network nation: human communication via computer.* Reading, MA: Addison-Wesley.

Hymes D (1986). 'Models of the interaction of language and social life.' In Gumperz J J & Hymes D (eds.) *Directions in sociolinguistics: the ethnography of speaking.* New York: Basil Blackwell. 35–71.

Kendall L (2002). *Hanging out in the virtual pub: masculinities and relationships online.* Berkeley: University of California.

Kiesler S, Siegel J & McGuire T W (1984). 'Social psychological aspect of computer-mediated communication.' *American Psychologist 39,* 1123–1134.

Koutsogiannis D & Mitsikopoulou B (2003). 'Greek and Greeklish.' *Journal of Computer-Mediated Communication 9(1).*

Lea M, O'Shea T, Fung P & Spears R (1992). '"Flaming" in computer-mediated communication: observations, explanations, implications.' In Lea M (ed.) *Contexts of Computer-Mediated Communication.* London: Harvester-Wheatsheaf. 89–112.

Ling R (2005). 'The socio-linguistics of SMS: an analysis of SMS use by a random sample of Norwegians.' In Ling R & Pederson P (eds.) *Mobile communications: renegotiation of the social sphere*. London: Springer. 335–349.

Mulkay M (1985). 'Agreement and disagreement in conversations and letters.' *Text 5,* 201–227.

Mulkay M (1986). 'Conversations and texts.' *Human Studies 9,* 303–321.

Myers D (1987). 'A new environment for communication play: online play.' In Fine G A (ed.) *Meaningful play, playful meaning*. Champaign, IL: Human Kinetics Publishers. 231–245.

Preece J & Ghozati K (1998). 'In search of empathy online: a review of 100 online communities.' In *Proceedings of the 1998 Association for Information Systems Americas Conference*. 92–94.

Rice R E & Love G (1987). 'Electronic emotion: socioemotional content in a computer-mediated communication network.' *Communication Research 14(1),* 85–108.

Savicki V, Lingenfelter D & Kelley M (1996). 'Gender language style and group composition in Internet discussion groups.' *Journal of Computer-Mediated Communication 2(3).*

Short J, Williams E & Christie B (1976). *The social psychology of telecommunications.* Chichester: Wiley.

Werry C C (1996). 'Linguistic and interactional features of Internet Relay Chat.' In Herring S (ed.) *Computer-mediated communication: linguistic, social, and cross-cultural perspectives.* Amsterdam: John Benjamins. 47–63.

Relevant Website

http://www.translate-to-success.com – Translate to Success.

Language in the Nondominant Hemisphere

S Campbell, The George Washington University, Alexandria, VA, USA

It is generally accepted that the right hemisphere makes significant contributions to language. Despite difficulties in research methodologies and interpretation of findings, particularly in regard to subject selection and population matching, viewed collectively the research findings from studies of both normal and brain-damaged populations provide a broad, but definitive, picture of the right hemisphere's contribution to language. In contrast to that of the left hemisphere, which is responsible for what is classically considered the linguistic attributes of language (the grammar, lexical semantics, and syntax), the right hemisphere's contributions to language are less constrained but "typically layered on top of the foundation provided by the left hemisphere" (Brownell, 2000). The right hemisphere makes significant contributions in the areas of prosody, the paralinguistic feature of a language commonly known as the 'melody' of speech; semantics; nuance and nonliteral language (abstract meanings); and discourse – in other words, those aspects of language linked to successful, interactive communication. After a brief discussion of the nature and history of the research, this article provides a broad overview of the language contributions of the right hemisphere to language, with particular emphasis on prosody, lexical semantics, and discourse. Finally, some comments regarding 'commonality' and directions for further research are presented.

History of Research and Comments

Early Research Thoughts and Shift in Direction

A review of the language contributions of the right hemisphere would not be complete without historical note of the early research in brain and language that focused primarily on the left hemisphere and its specialization and dominance for language.

The framework for early research was defined largely by Paul Broca (1865), who declared the left hemisphere (LH) dominant for language based on his research with brain-damaged patients who experienced aphasia secondary to lesions in the left frontal lobe. Driven by this assertion of LH dominance for language, the quest to define language abilities neurolinguistically essentially became a 'left hemisphere task.' As a result, early research focused on localization of specific function within the dominant LH. The primary means of investigating dominance and localization of function was through clinical observations that compared differences in language disturbances between patients with extensive left and right hemisphere damage. Observations consistently indicated that patients with left hemisphere damage (LHD) presented obvious language deficits that were not observed in patients with similarly extensive right hemisphere damage (RHD). Despite this 'left bias,'

Lea M, O'Shea T, Fung P & Spears R (1992). '"Flaming" in computer-mediated communication: observations, explanations, implications.' In Lea M (ed.) *Contexts of Computer-Mediated Communication.* London: Harvester-Wheatsheaf. 89–112.
Ling R (2005). 'The socio-linguistics of SMS: an analysis of SMS use by a random sample of Norwegians.' In Ling R & Pederson P (eds.) *Mobile communications: renegotiation of the social sphere*. London: Springer. 335–349.
Mulkay M (1985). 'Agreement and disagreement in conversations and letters.' *Text 5,* 201–227.
Mulkay M (1986). 'Conversations and texts.' *Human Studies 9,* 303–321.
Myers D (1987). 'A new environment for communication play: online play.' In Fine G A (ed.) *Meaningful play, playful meaning*. Champaign, IL: Human Kinetics Publishers. 231–245.
Preece J & Ghozati K (1998). 'In search of empathy online: a review of 100 online communities.' In *Proceedings of the 1998 Association for Information Systems Americas Conference*. 92–94.
Rice R E & Love G (1987). 'Electronic emotion: socioemotional content in a computer-mediated communication network.' *Communication Research 14(1),* 85–108.
Savicki V, Lingenfelter D & Kelley M (1996). 'Gender language style and group composition in Internet discussion groups.' *Journal of Computer-Mediated Communication 2(3).*
Short J, Williams E & Christie B (1976). *The social psychology of telecommunications.* Chichester: Wiley.
Werry C C (1996). 'Linguistic and interactional features of Internet Relay Chat.' In Herring S (ed.) *Computer-mediated communication: linguistic, social, and cross-cultural perspectives.* Amsterdam: John Benjamins. 47–63.

Relevant Website

http://www.translate-to-success.com – Translate to Success.

Language in the Nondominant Hemisphere

S Campbell, The George Washington University, Alexandria, VA, USA

It is generally accepted that the right hemisphere makes significant contributions to language. Despite difficulties in research methodologies and interpretation of findings, particularly in regard to subject selection and population matching, viewed collectively the research findings from studies of both normal and brain-damaged populations provide a broad, but definitive, picture of the right hemisphere's contribution to language. In contrast to that of the left hemisphere, which is responsible for what is classically considered the linguistic attributes of language (the grammar, lexical semantics, and syntax), the right hemisphere's contributions to language are less constrained but "typically layered on top of the foundation provided by the left hemisphere" (Brownell, 2000). The right hemisphere makes significant contributions in the areas of prosody, the paralinguistic feature of a language commonly known as the 'melody' of speech; semantics; nuance and nonliteral language (abstract meanings); and discourse – in other words, those aspects of language linked to successful, interactive communication. After a brief discussion of the nature and history of the research, this article provides a broad overview of the language contributions of the right hemisphere to language, with particular emphasis on prosody, lexical semantics, and discourse. Finally, some comments regarding 'commonality' and directions for further research are presented.

History of Research and Comments

Early Research Thoughts and Shift in Direction

A review of the language contributions of the right hemisphere would not be complete without historical note of the early research in brain and language that focused primarily on the left hemisphere and its specialization and dominance for language.

The framework for early research was defined largely by Paul Broca (1865), who declared the left hemisphere (LH) dominant for language based on his research with brain-damaged patients who experienced aphasia secondary to lesions in the left frontal lobe. Driven by this assertion of LH dominance for language, the quest to define language abilities neurolinguistically essentially became a 'left hemisphere task.' As a result, early research focused on localization of specific function within the dominant LH. The primary means of investigating dominance and localization of function was through clinical observations that compared differences in language disturbances between patients with extensive left and right hemisphere damage. Observations consistently indicated that patients with left hemisphere damage (LHD) presented obvious language deficits that were not observed in patients with similarly extensive right hemisphere damage (RHD). Despite this 'left bias,'

the RH's capabilities in the area of language did not go unnoticed. For example, Karl Wernicke (1874) suggested that the right, or nondominant, hemisphere made contributions to language. He based his speculations on the recovery of language functions in the presence of LH damage, proposing that the undamaged RH was responsible for assuming the recovered language functions. Unfortunately, since the clinical observations that dominated the research did not support theories of RH involvement, research with brain-damaged populations continued to focus on language as the domain of the LH. Then, in the 1960s, the so-called 'silent' RH began to receive attention when commissurotomies allowed the two hemispheres to be studied independently. Although much of the focus of this early 'split brain' research was in the area of visual perception, attention was focused on suggestions regarding the 'isolated' RH's ability to decode simple linguistic forms in the absence of interhemispheric communication (LH input) via the corpus callosum. In addition, at approximately the same time, a study involving RHD patients conducted by Eisenson (1962) suggested RH involvement in language processing. Eisenson evaluated the recognition vocabulary and sentence processing of grammatical correctness in 65 RHD patients. Based on his subjects' performance on these tasks, Eisenson argued that RHD damaged patients were deficient in both vocabulary and sentence processing. Although the research model and procedures of this study were viewed by some as flawed, these results, along with those of the split brain studies, supported previously held theoretical assumptions and served to initiate serious research interest regarding the language contributions of the RH.

In the early 1970s, interest in the language contributions of the nondominant or right hemisphere increased dramatically. Researchers working with brain-damaged populations began to explore the nature of the RH contribution to language. The introduction of noninvasive laterality techniques (dichotic listening and visual tachistoscopy) that allowed investigators to study the contributions of each hemisphere to a variety of tasks in nonbrain-damaged populations added another dimension to the study of RH language. Occurring at approximately the same time was a new focus on language and communication within a pragmatic context and an increase in the understanding of brain–behavior relationships throughout development. Both the increased knowledge regarding development and an additional framework for studying language were also critical to the growing research interest in RH language functions. Finally, research in cognitive neurosciences pointed to apparent processing differences in the left and right hemispheres that held implications for hemispheric differences in language. Functional dichotomies (left brain–right brain), also known as 'learning styles,' were established to depict the specialized processing styles of each hemisphere. The left, or verbal, hemisphere was believed to process highly structured information in a logical, step-by-step, analytical manner using verbal symbols such as words. In contrast, the right, visual or nonverbal, hemisphere processed information in a holistic, visual spatial manner using pictures and images rather than words. Evaluating hemispheric contributions to language in terms of each hemisphere's dominant processing style strengthened arguments for a RH contribution in various language tasks associated with a particular processing style (e.g., holistic and comprehension of main idea) and added another framework for the research – one defined by processing styles.

Critical Evaluation

The research that followed the interest generated in the 1970s utilized both the methodologies and the techniques critical to our understanding of LH language functioning (the presence of damage to one or both of the hemispheres) as well as newly developed methodologies and laterality techniques (e.g., dichotic listening) that enabled the study of hemispheric specialization of language abilities in normal populations. In addition, research questions expanded as RH language contributions were studied within the frameworks of cognitive processing styles and pragmatics. The result was the development of a rich body of evidence from widely diverse populations using a number of different methods of study. This evidence not only confirmed that the RH made significant contributions to language but also shed light on the nature of the contributions.

Although the resulting 'picture' of RH language capability and contribution is generally accepted, inconsistencies in results and interpretations frequently occur. Much of this can be explained by virtue of the complex nature of the subject of study and the diverse experimental designs and varying theoretical models employed. However, despite numerous arguments pro and con regarding the 'best' method of study or population (clinical observation vs. laterality techniques; HD vs. normal or adult vs. child), it is important to remember that each study must be evaluated critically due to problems inherent in brain behavior research. For example, studies of brain-damaged populations are generally, and correctly, considered to provide the most definitive information regarding hemispheric specialization and contribution to language. However, the RH contributions to

language are defined in terms of that performance resulting from damage to either the LH or the RH. The resulting language sequelae are assumed to allow one to infer and assign RH capabilities based on the presence or absence of language behavior relative to hemispheric damage. Care to define and match the populations becomes a difficult and critical issue, particularly since localization of damage is not necessarily a consistent, stable predictor of actual performance, except, perhaps, in the case of developmental hemispherectomy (Dennis and Whitaker, 1977). Studies of RH language involving normal, non-brain-damaged populations have generally relied on laterality techniques such as dichotic listening, tachistoscopic visual presentations, sodium amytal (Wada test), regional blood flow, and, recently, sophisticated neuroimaging. These methods also deserve critical scrutiny when interpreting results. In normal, nonbrain-damaged populations, no matter what the investigative technique, it is difficult, if not impossible, to completely isolate the left and right cerebral hemispheres due to the input received via the corpus callosum. In addition, there is also strong evidence of involvement of subcortical structures in processing language, which further complicates the 'isolation' of one hemisphere from the other. Finally, a limited number of studies have examined the RH contributions to language in children diagnosed with nonverbal learning disabilities. The nonverbal learning disability (NLD) is defined primarily by significant gaps in verbal and nonverbal intellectual abilities or intelligence quotients on the Wechsler Intelligence Scale for Children, with nonverbal performance abilities (consisting of tasks primarily mediated by the RH) significantly lower than verbal abilities (tasks primarily mediated by the LH) (Rourke, 1989). These studies usually approach the task of assigning language attributes from a 'learning styles' model that tends to address the language issues in a less specific manner than does the body of research with brain-damaged populations. However, the RH language profile in the presence of NLD is strikingly similar to that found in more 'classical' methods of investigation. The difficulties in interpretation with such an approach are obvious, as are those encountered when comparing adult and child populations, a developed versus a developing language system.

In summary, despite the need for prudence regarding interpretations of research methodologies and findings, the conclusions drawn from the vast majority of studies of brain language relationships not only support the notion of the RH contribution to language but also are remarkably consistent in identifying the salient features of the contribution and the disrupted functions that arise in the presence of RH damage or related cognitive disabilities.

Right Hemisphere Language Contributions

A compilation of the data regarding language capabilities of the RH reveals a remarkably diverse range of contributions in the areas of reading comprehension; auditory comprehension; prosodic comprehension and expression; lexical semantics; the comprehension of abstract or nonliteral meanings such as metaphor, proverbs, and figurative language; the comprehension of main ideas; and the comprehension and organization of narratives and the rules of discourse. Simply considered, disruption of input from the RH produces deficits that mirror the contributions. For example, some RHD patients may present as overly literal, rambling, too focused on detail, and unable to 'make a point' (lack of main idea).

The contributions of the RH, whether defined in terms of their absence (language deficits) or performance, are evidenced across each of the four modalities of language, speaking, listening, reading, and writing. The specific contribution is made as a primarily expressive language function, a primarily receptive language function, or a contribution that has equitable dual representation in the comprehension and production of language. The impact on communicative competence varies with the degree of RHD or, in the case of NLD, the collective absence of RH language contributions in conjunction with nonverbal cognitive deficits. Consequently, the level of disruption of communicative functioning has a broad range, from subtle deficits with minor impacts on communication and learning to profound deficits such as the so-called 'language without communication' seen in Asperger's syndrome or deep dyslexia.

Prosody

Prosody, the melodic line of speech produced by variations in fundamental frequency, intensity (loudness), and duration (length), provides meaning through paralinguistic cues and serves as an organizational framework for language that also plays a major role in the language acquisition process. Within investigative models, prosodic contributions to language meaning are delineated into four distinct types based on the paralinguistic meaning conveyed: (1) linguistic or intrinsic prosody, which signals questions, statements, etc., or distinguishes meanings of words and phrases ('yellow jacket' meaning a bee

or an article of clothing); (2) affective or emotional prosody, which signals anger, pleasure, or other emotional states; (3) intellectual prosody, which signals nuances such as sarcasm; and (4) inarticulate prosody, which signals agreement through the use of 'intellectual grunts' such as 'uh huh' (Monrad-Krohn, 1947). The majority of the research in hemispheric participation in the processing of prosody has focused primarily on the first two types of prosody, linguistic and affective prosody. In general, the results of numerous studies on brain-damaged populations indicate a role for the RH in the production and comprehension of both linguistic and affective prosody, with the most marked disturbances occurring in affective prosody.

Studies of Affective Prosody

The majority of the investigations of affective prosody indicate a RH dominance for the production and comprehension of affective prosody in brain-damaged populations (Benowitz *et al.*, 1983; Code, 1987). In addition, a few studies involving the processing of affective prosody by normal subjects using dichotic listening paradigms offer further support for the RH's prosodic language capabilities because results support RH participation in processing affective prosody (e.g., Herrero and Hillix, 1990). From this evidence and that of other similar studies, a profile of affective prosodic disturbances emerges. In the presence of RHD, production or the ability to convey emotional prosody can be characterized as flat or monotonous with obvious timing disturbances, expressed as 'abnormally uniform' pauses. In contrast, the comprehension deficits of affective prosodic disturbances are expressed by both normal subjects and those with RHD as difficulty in perceiving pitch changes and a "reduced ability to discriminate and identify mood in neutral sentences and sentences with incongruent meaning and prosodic contour" (Meyers, 1999, p. 82).

Although there is strong evidence of a RH dominance for affective prosody, the level of involvement and specific nature of the contribution of the RH in the production and processing of linguistic prosody are less definitive. The results of research involving processing and production of linguistic prosody by normal subjects and RHD patients are somewhat less consistent, with a few studies suggesting bilateral involvement in processing linguistic prosody or a RH dominance for affective prosody only. The inconsistent findings and sometimes conflicting results may be due, in part, to the multiple parameters involved in conveying linguistic information through prosody – the use of tone, stress, duration, and intonation – to communicate linguistic information (Hird and Kirsner, 1993). For example, in a study involving contrastive stress, a type of linguistic prosody, (Weintraub *et al.*, 1981) found that RHD subjects performed more poorly than controls on tasks involving the production and perception of this measure of 'nonemotional' prosody. In contrast, results from studies that evaluated linguistic prosody within the parameter of pitch or tone signaling linguistic meaning at the lexical level failed to find evidence to support RH involvement. Regarding the inconsistencies, the results often appear to vary with respect to the parameter of linguistic prosody being evaluated. The exception appears to be the parameter of intonation, which seems to offer consistent evidence of RH participation (Grant and Dingwall, 1984).

Despite the inconsistencies mentioned previously, a number of studies on RHD patients have offered clear evidence of RH participation in processing linguistic prosody, with both expressive and receptive deficits occurring in the presence of damage. From these studies, a profile of prosodic deficits occurring in RHD likewise emerges. For RHD patients, deficits in the production of linguistic prosody are expressed as a difficulty modulating pitch and limited use of word stress to signal meaning in sentences. Similarly, the comprehension deficits also involve stress and pitch and difficulty discerning meaning at both the word and the sentence level based on stress patterns, difficulty discerning sentence type from prosodic contour (Weintraub *et al.*, 1981), and an insensitivity of prosodic cues (Baum and Dwivedi, 2003).

Lexical Semantics

The body of research in the area of semantics and RH processing and production is smaller and less definitive than that involving prosody. Part of the reason is due to the general difficulties encountered in studying the linguistic system of language, which determines meaning and encompasses the relationships between language, thought, and behavior. Accordingly, research findings of this complex language behavior tend to be best understood when processing and production are considered 'holistically' as dependent functions, not easily separated into production–comprehension dichotomies. At best, findings in semantic production tasks (e.g., naming and word retrieval) are inconsistent and appear to be highly task dependent, suggesting that the 'production' deficits may not be a primary deficit but, rather, one secondary to comprehension (processing) deficits (Gainotti *et al.*, 1983).

Investigations of semantic processing on both normal and brain-damaged populations indicate

qualitative differences between the two hemispheres in connotative and denotative semantic associations (Zaidel, 1985). In general, the RH appears sensitive to connotative associations – those associations that are remote and implicit and persist over time and adept at maintaining multiple meanings (Beeman, 1998). The RH is involved in semantic tasks that entail wide-ranging response sets, which include nondominant, less familiar, connotative meanings and words related by category rather than associative function (Chiarello *et al*., 1990). In addition, a number of investigations suggest that the RH semantic processing occurs at multiple levels, and that the level of RH involvement is specific to the type of semantic information being processed and the level of association (e.g., concrete vs. abstract nouns; literal or close vs. 'loose' or distant associations) (Beeman, 1998; Seger *et al*., 2000). Finally, in addition to semantic contributions at the word level, there is evidence of RH involvement in lexical semantics at the phrase and sentence level of interpretation in metaphor (Van Lancker and Kempler, 1987) and in nuance (Rodel *et al*., 1992), semantic tasks that rely on 'distant' meaning relationships and the ability to resolve lexical ambiguity.

Individuals with RHD tend to be overly literal and fail to recognize 'shades' of differences in meaning at the word level (Brownell *et al*., 1984). There is further evidence indicating that the RHD patients' tendency toward 'literalness' has disrupted semantic relationships (Chiarello and Church, 1986) and semantic processes in lexical judgment tasks involving single word semantic similarity. RHD subjects do not seem to have difficulty in comprehending the dominant meanings of single words (a LH function); however, divergent processing tasks that demand sensitivity to relationships and broad meanings present considerable difficulty relative to the degree of semantic overlap (Beeman, 1998; Brownell *et al*., 1984; Chiarello *et al*., 1990). Considered collectively, studies of lexical semantics in RHD subjects indicate a lack of semantic organizing principles that are critical in lexical semantic processing and performance.

Discourse and Pragmatics

RH deficits compromise complex language functions in a nonspecific manner that belies simple explanations (McDonald, 2000). This statement, although relevant to the study of language in general, is particularly true when one attempts to clarify issues regarding the RH and discourse and pragmatics. The two terms generally overlap but convey slightly different meanings depending on the author describing the terms. For the purposes of this article and in an effort to add clarity, discourse will be considered as a pragmatic function specifically concerned with cohesion and coherence (integrating meaning) and deriving the general idea from the whole message through complex and abstract linguistic interpretation. Thus defined, discourse is not limited to conversation but has various other forms – procedural, narrative, and expository (Hough and Pierce, 1993). Also, as in semantics, production and comprehension discourse capabilities are difficult to isolate and are best understood holistically, outside the usual dichotomous model of comprehension versus production.

Pragmatics can be thought of as the 'hows' of communication – how and how well we make our point. By definition, it encompasses the rules underlying conversational discourse, the function of language particular to participants, situational knowledge, the topic at hand, related 'world' knowledge or information, and social conventions. It includes both the context and the intent of a message and when applied to discourse includes additional rules to enable interactive communication (turn-taking, topic maintenance, etc.). Pragmatics integrates language functions mediated by the two hemispheres with major contributions from the RH in two areas discussed previously (prosody and semantics) and also refers to the nonverbal skills used in communication and the communication-related disturbances believed to be mediated by the RH.

Discourse deficits are the hallmark of the communication difficulties of RHD patients (Meyers, 1999). Results of numerous studies on RHD patients indicate RH involvement in discourse (integrating meaning and determining the general idea from the whole message). The characteristics of the discourse disturbances in the presence of RHD are wide ranging, with particular difficulty in the comprehension and production of narratives, generally considered the domain of the RH. In complex or lengthy narratives, the absence of a unifying theme or main idea causes particular problems in comprehension and retention. In discourse interpretation, main ideas serve to provide structure for identifying and organizing key concepts rather than the 'details' underlying those concepts, thus enhancing comprehension and recall. In support of this notion, a number of studies suggest that basic discourse form is retained within structured tasks (e.g., providing the order of a short story when given individual sentences presented in random order (Huber and Gleber, 1982), predicting story endings (Rehak *et al*., 1992), and drawing simple inferences in short stories and interpreting sarcasm in comments in forced choice format (Kaplan *et al*., 1990)). In contrast, it has been suggested that in RHD the comprehension of complex narratives poses particular

problems that may be related to the need for reinterpretation (Brownell *et al.*, 2000), a task that relies heavily on the identification of the main idea to define key concepts and aid recall. RHD subjects also have difficulty recognizing purpose or intent of the narrative (a skill related to comprehension of theme and main idea), drawing inferences, inferring motives in complex narratives (Wapner *et al.*, 1981), predicting outcomes in narratives without explicit cues (Rehak *et al.*, 1992), applying alternate meanings, and understanding implied and nonliteral meanings. In discourse production, RHD subjects have difficulty following conversational rules and a propensity for excessive detail with an absence of a unifying theme or idea. Similar problems occur in writing, which is also characterized by excess detail and an absence of unifying main ideas, transitions, and conclusions. Finally, with respect to pragmatic functions during conversational discourse, RHD patients have difficulty interpreting speaker intent and indirect requests and evaluating paralinguistic cues.

Concluding Comments and Directions for Research

The range of language and communication skills associated with the RH is extensive. This article has attempted to present a wide-ranging overview of the attributes of RH language capabilities identified during the past 25 years and to focus on three areas of language in which the RH role is significant: prosody, lexical semantics, and discourse and pragmatics. Although some contradictions and inconsistencies can be found when comparing studies of RH language, the presence of unique RH participation in language is generally accepted. The general profiles of contributions that emerge are strikingly similar despite differences in investigative populations (brain-damaged, normal, child, or adult) and methodologies (studying clinical manifestations of damage, normal participation in lateralization studies, or describing abilities within learning disabled populations, specifically NLD). Some confusion remains when attempting to define deficit profiles, primarily because RHD or disruption of RH language does not necessarily impact all language modalities simultaneously. For example, one may have severe deficits in reading comprehension that are not evident in the comprehension of spoken language. Furthermore, the severity of the disturbances is not systematically expressed. Severity can vary across all modalities and can range from severe to subtle dependent on the task, language system (prosody, lexical semantics, etc.), or modality. We need to strengthen our understanding of the attributes of RH language contributions and the language impacts of RHD or impaired RH language function. Valuable information can be obtained from two sources – investigative research and clinical data.

Traditional Research Models

Traditional research models have provided an extensive picture of RH language contributions in both normal and RHD populations. In-depth studies need to be performed that narrow the focus and investigate specific aspects of RH language 'abilities' or deficits (e.g., main idea vs. the broader domain of discourse), with particular attention to differences in performance across modalities. Future research should include ways to objectively define the degree of impact (subtle to profound) that the RHD or RH deficits produce relative to particular language tasks or within language domains. The confounding affects of modality and the level of functional impact of RH disturbance (subtle to profound) could then be addressed or controlled in future research and possibly clarify some of the inconsistencies across studies. Research attention should also be given to localization of function, which is poorly understood. In addition to the difficulty specifying location of function, there is evidence of involvement of subcortical structures in language production and processing. Studies that take advantage of advances in brain imaging technology should provide clarity, but currently localization of function remains unclear. Finally, a number of studies have provided evidence that the RH is capable of mediating language if the LH is damaged early in life (Dennis and Whitaker, 1977; Reilly *et al.*, 2004); however, we need to know more about the nature of the RH language contributions in developing neurological systems. Investigations should include normally developing and learning disabled (NLD) children as well as children with brain damage.

Data from Clinical Interventions

Our knowledge of the RH contributions to language can also be enhanced by carefully documenting the outcome of therapeutic interventions with RHD populations and language disordered children who exhibit language deficits generally associated with RHD. Relatively little has been published regarding the outcome of therapeutic intervention in disordered populations with RHD or 'RH language disorders.' Current knowledge regarding the contributions of the RH to language processing and production should allow the development of effective remediation models designed to address objectively defined RH deficits in a structured and systematic fashion. Documentation of the course of recovery of function in response to such intervention strategies would not only assist

in the development of more effective remediation models but also provide information that has special relevance for RH language research. Major questions need to be answered regarding the type and degree of deficit relative to prognosis and disparate responses to intervention. For example, is prognosis for recovery relative only to degree or severity of impact or also relative to particular language domains (e.g., are semantic deficits more difficult to resolve that written narrative organization)? How are prognosis and/or recovery of function related to the LH's ability to subsume a RH language function? In other words, what can we infer regarding RH abilities when recovery takes place? Where is recovery localized? These are just a few of the questions that carefully controlled investigations of recovery of function relative to remediation could address. Clearly, integrating the findings from traditional research studies and the data resulting from therapeutic intervention holds promise for addressing the many unresolved issues regarding language and the nondominant hemisphere.

See also: Anatomical Asymmetries versus Variability of Language Areas of the Brain; Laterality Across the World's Languages.

Bibliography

Baron-Cohen S, Tager-Flusherg H & Cohen D J (eds.) (2000). *Understanding other minds: perspectives from autism and developmental cognitive neuroscience* (2nd edn.). Oxford: Oxford University Press.

Baum S & Dwivedi V (2003). 'Sensitivity to prosodic structure in left- and right-hemisphere-damaged individuals.' *Brain and Language 87,* 278–289.

Beeman M (1998). 'Coarse semantic coding and discourse comprehension.' In Beeman M & Chiarello C (eds.) *Right hemisphere language comprehension: perspectives from cognitive neuroscience.* Mahwah, NJ: Erlbaum. 255–284.

Benowitz L I, Bear D M, Rosenthal R, Mesulam M, Zaidel E & Sperry R W (1983). 'Hemispheric specialization in non-verbal communication.' *Cortex 19,* 5–14.

Broca P (1865). 'Sur la siège de la faculté du language articulé.' *Bulletins de la Société d'Anthropologie (Paris) 6,* 337–393.

Brownell H (2000). 'Right hemisphere contributions to understanding lexical connotation and metaphor.' In Grodzinsky Y, Shapiro L & Swinney D (eds.) *Language and the brain.* San Diego: Academic Press. 185–201.

Brownell H, Griffin R, Winner E *et al.* (2000). 'Cerebral lateralization and theory of mind.' In Baron-Cohen S, Tager-Flushberg H & Cohen D J (eds.). 311–338.

Brownell H H, Potter H H & Michelow D (1984). 'Sensitivity to lexical denotation and connotation in brain-damaged patients: a double dissociation?' *Brain and Language 22,* 253–265.

Chiarello C, Burgess C, Richards L *et al.* (1990). 'Semantic and associative priming in the cerebral hemispheres: some words do, some words don't ... sometimes, some places.' *Brain and Language 38,* 75–104.

Chiarello C & Church K L (1986). 'Lexical judgments after right or left hemisphere injury.' *Neuropsychologia 24,* 623–640.

Code C (1987). *Language aphasia in the right hemisphere.* New York: John Wiley and Sons.

Dennis M & Whitaker H A (1977). 'Hemispheric equipotentiality and language acquisition.' In Segalowitz S F & Gruber F (eds.) *Language development and neurological theory.* New York: Academic Press. 94–104.

Eisenson J (1962). 'Language and intellectual modifications associated with right cerebral damage.' *Language and Speech 5,* 49–53.

Gainotti G, Caltagrione C & Miceli G (1983). 'Selective impairment of semantic-lexical discrimination in right-brain-damaged patients.' In Perecman E (ed.) *Cognitive processing in the right hemisphere.* New York: Academic Press. 149–167.

Grant S R & Dingwall W O (1984). The role of the right hemisphere in processing linguistic prosody. Presentation at the Academy of Aphasia, October 1984, Los Angeles, California.

Herrero J V & Hillix W A (1990). 'Hemispheric performance in detecting prosody: a competitive dichotic listening task.' *Perceptual and Motor Skills 71,* 470–486.

Hird K & Kirsner K (1993). 'Dysprosody following acquired neurogenic impairment.' *Brain and Language 45,* 46–60.

Hough M & Pierce R S (1993). 'Contextual and thematic influences on narrative comprehension of left and right hemisphere brain damaged patients.' In Brownell H H & Joanette Y (eds.) *Narrative discourse and neurologically impaired and normal aging adults.* San Diego: Singular. 213–238.

Huber W & Gleber J (1982). 'Linguistic and nonlinguistic processing of narratives in aphasia.' *Brain and Language 16,* 1–18.

Joanette Y, Goulet P & Hannequin D (1990). *Right hemisphere and verbal communication.* New York: Springer-Verlag.

Kaplan J A, Brownell H H, Jacobs J R *et al.* (1990). 'The effects of right hemisphere damage on the pragmatic interpretation of conversational remarks.' *Brain and Language 38,* 315–333.

McDonald S (2000). 'Exploring the cognitive basis of right-hemisphere pragmatic language disorders.' *Brain and Language 75,* 82–107.

Meyers P (1999). *Right hemisphere damage: disorders of communication and cognition.* San Diego: Singular.

Monrad-Krohn G H (1947). 'Dysprosody or altered "melody of language."' *Brain 70,* 405–415.

Rehak A, Kaplan J A, Weylman S T *et al.* (1992). 'Story processing in right-hemisphere brain-damaged patients.' *Brain and Language 42,* 320–336.

Reilly J, Losh M, Bellugi U *et al.* (2004). '"Frog, where are you?" Narratives in children with language impairment,

early focal brain injury, and Williams syndrome.' *Brain and Language 88,* 229–247.
Rodel M, Cook N & Regard M (1992). 'Hemispheric dissociation in judging semantic relations: complementarity for close and distant associations.' *Brain and Language 43,* 448–459.
Rourke B P (1989). *Nonverbal learning disabilities: the syndrome and the model.* New York: Guilford.
Seger C A, Desmond J E, Glover G *et al.* (2000). 'Functional magnetic resonance imaging evidence for right-hemisphere involvement in processing unusual semantic relationships.' *Neuropsychology 4(3),* 361–369.
Stemmer B & Whitaker H A (eds.) (1998). *Handbook of neurolinguistics.* San Diego: Academic Press.
Van Lancker D R & Kempler D (1987). 'Comprehension of familiar phrases by left but not by right hemisphere damaged patients.' *Brain and Language 32,* 265–277.
Wapner W, Hampy S & Gardner H (1981). 'The role of the right hemisphere in the appreciation of complex linguistic materials.' *Brain and Language 14,* 15–33.
Weintraub S, Mesulam M & Kramer L (1981). 'Disturbances in prosody – a right hemisphere contribution to language.' *Archives of Neurology 38,* 742–744.
Wernicke C (1874). *Der aphasische symptomenkomplex.* Breslau, Germany: Cohn Weigert.
Zaidel E (1985). 'Language in the right hemisphere.' In Benson D F & Zaidel E (eds.) *The dual brain.* New York: Guilford. 205–231.

Language in the Workplace: Different Approaches

F Bargiela-Chiappini, Nottingham Trent University, Nottingham, UK

The selective chronology adopted in this discussion documents changes in the nature and scope of the field since its inception in the 1960s. In the early days, LSP (Language for Specific Purposes) and later ESP (English for Specific Purposes) dominated the scene and have continued to attract a steady following alongside the other vocational stream of language needs analysis. More recently, discourse analytic and ethnographic studies relying on empirical data and an ever increasing interest in intercultural business interactions (*see* **Intercultural Pragmatics and Communication**), especially negotiations and meetings, have enriched and expanded the original vocational thrust of the field.

The dynamics of change within Language of Work (LoW) over the last two decades have fundamentally and, arguably, permanently changed its nature while continuously challenging its established boundaries and traditional concerns. In the 1980s, the 'linguistic turn' caused a reappraisal of the value and possibilities afforded by internal or cross-company quantitative surveys or microanalyses of isolated text samples. Increasingly, situated pragmatic and sociolinguistic approaches to the language of work and a discursive approach to its social-psychological dimensions are opening up new prospects for multimethod research and crossdisciplinary dialogue.

Furthermore, advances in communication technology also have altered the workplace beyond recognition, affecting interpersonal and intercompany communication and have contributed to the internationalization of cultural and economic exchange, or globalization. 'Localization,' a centripetal movement contrasting globalization and its contested imposition of 'Americanization' (*see* **Linguistic Imperialism**), is leading to the rediscovery of the importance and the role of other languages and value systems and is challenging the long-established position of English as a *lingua franca* on the scene of international business.

Naturally, macrovariation in LoW research has also affected topic choice and methodological approaches at the microanalytical level. For instance, language needs analysis with its impersonal and quantitative understanding of language(s) in the workplace has been largely replaced by the systematic analysis of situated language practices conducted by scholars from a variety of disciplines and using multimethod research, increasingly in collaboration with the actual practitioners (*see* **Language in the Workplace**).

From Language to Discourse

In its early days, concern for users' needs in industry had led to the formation of a strong body of research on Language for Specific Purposes (LSP), which concentrated on the analysis and composition of practical text-types as required in the workplace (*see* **Language for specific purposes**). With the emergence of English as a *lingua franca*, English for Specific Purposes (ESP) came into existence to undertake the analysis of the linguistic features of specialist texts, at first through register analysis (*see* **Registers and language varieties in cultures**) and later through genre analysis (Dudley-Evans and St John, 1998).

The late 1980s and early 1990s witnessed the emergence within the LoW of a strong body of discourse-based scholarship, where discourse was variously understood to refer to, at one end of the spectrum, the

early focal brain injury, and Williams syndrome.' *Brain and Language 88,* 229–247.

Rodel M, Cook N & Regard M (1992). 'Hemispheric dissociation in judging semantic relations: complementarity for close and distant associations.' *Brain and Language 43,* 448–459.

Rourke B P (1989). *Nonverbal learning disabilities: the syndrome and the model.* New York: Guilford.

Seger C A, Desmond J E, Glover G *et al.* (2000). 'Functional magnetic resonance imaging evidence for right-hemisphere involvement in processing unusual semantic relationships.' *Neuropsychology 4(3),* 361–369.

Stemmer B & Whitaker H A (eds.) (1998). *Handbook of neurolinguistics.* San Diego: Academic Press.

Van Lancker D R & Kempler D (1987). 'Comprehension of familiar phrases by left but not by right hemisphere damaged patients.' *Brain and Language 32,* 265–277.

Wapner W, Hampy S & Gardner H (1981). 'The role of the right hemisphere in the appreciation of complex linguistic materials.' *Brain and Language 14,* 15–33.

Weintraub S, Mesulam M & Kramer L (1981). 'Disturbances in prosody – a right hemisphere contribution to language.' *Archives of Neurology 38,* 742–744.

Wernicke C (1874). *Der aphasische symptomenkomplex.* Breslau, Germany: Cohn Weigert.

Zaidel E (1985). 'Language in the right hemisphere.' In Benson D F & Zaidel E (eds.) *The dual brain.* New York: Guilford. 205–231.

Language in the Workplace: Different Approaches

F Bargiela-Chiappini, Nottingham Trent University, Nottingham, UK

The selective chronology adopted in this discussion documents changes in the nature and scope of the field since its inception in the 1960s. In the early days, LSP (Language for Specific Purposes) and later ESP (English for Specific Purposes) dominated the scene and have continued to attract a steady following alongside the other vocational stream of language needs analysis. More recently, discourse analytic and ethnographic studies relying on empirical data and an ever increasing interest in intercultural business interactions (*see* **Intercultural Pragmatics and Communication**), especially negotiations and meetings, have enriched and expanded the original vocational thrust of the field.

The dynamics of change within Language of Work (LoW) over the last two decades have fundamentally and, arguably, permanently changed its nature while continuously challenging its established boundaries and traditional concerns. In the 1980s, the 'linguistic turn' caused a reappraisal of the value and possibilities afforded by internal or cross-company quantitative surveys or microanalyses of isolated text samples. Increasingly, situated pragmatic and sociolinguistic approaches to the language of work and a discursive approach to its social-psychological dimensions are opening up new prospects for multimethod research and crossdisciplinary dialogue.

Furthermore, advances in communication technology also have altered the workplace beyond recognition, affecting interpersonal and intercompany communication and have contributed to the internationalization of cultural and economic exchange, or globalization. 'Localization,' a centripetal movement contrasting globalization and its contested imposition of 'Americanization' (*see* **Linguistic Imperialism**), is leading to the rediscovery of the importance and the role of other languages and value systems and is challenging the long-established position of English as a *lingua franca* on the scene of international business.

Naturally, macrovariation in LoW research has also affected topic choice and methodological approaches at the microanalytical level. For instance, language needs analysis with its impersonal and quantitative understanding of language(s) in the workplace has been largely replaced by the systematic analysis of situated language practices conducted by scholars from a variety of disciplines and using multimethod research, increasingly in collaboration with the actual practitioners (*see* **Language in the Workplace**).

From Language to Discourse

In its early days, concern for users' needs in industry had led to the formation of a strong body of research on Language for Specific Purposes (LSP), which concentrated on the analysis and composition of practical text-types as required in the workplace (*see* **Language for specific purposes**). With the emergence of English as a *lingua franca*, English for Specific Purposes (ESP) came into existence to undertake the analysis of the linguistic features of specialist texts, at first through register analysis (*see* **Registers and language varieties in cultures**) and later through genre analysis (Dudley-Evans and St John, 1998).

The late 1980s and early 1990s witnessed the emergence within the LoW of a strong body of discourse-based scholarship, where discourse was variously understood to refer to, at one end of the spectrum, the

situated analysis of interactional exchanges; at the other end, the ways of conceptualizing and interpreting work practices within organizations. This development has generated a number of classics, which exemplify its methodological diversity: so, for example, Drew and Heritage (1992) and Boden (1994) apply conversation analysis (*see* **Conversation Analysis**), while Firth (1995) and Ehlich and Wagner (1995) concentrate mainly on discourse analytical studies.

It is also interesting to note the terminological evolution that has accompanied discourse-based research: from Institutional Discourse (Agar, 1985), to Professional Discourse (Gunnarsson, *et al.,* 1997), which embraces legal, medical, social welfare, educational, and scientific fields (*see* **Institutional Talk**), to Business Discourse (Bargiela-Chiappini and Harris, 1997a). In both former types of discourse, one of the interactants is usually a lay person; in contrast, business discourse is seen to take place between individuals whose main activities are located within business and whose interactions are motivated by matters relating to their respective businesses.

Worth highlighting is the distinctive contribution to the LoW from North America. Management communication, technical communication, English composition, and business communication, the latter characterized by a strong vocational orientation, are all fields with a distinctive understanding of communication and language, a more recent complementary field being organizational communication (Jablin and Putman, 2001). Discussion of the potential contribution of these fields to a multidisciplinary scholarly base for the LoW has been initiated, which seeks to involve linguists and sociolinguists in Europe and beyond (Bargiela-Chiappini and Nickerson, 2001).

The LoW has recently benefited from sustained efforts by applied linguists to engage with issues of relevance in research practice (Sarangi and Candlin, 2003). Accordingly, academic researchers are called on to collaborate with practitioners in the task of formulating a joint understanding of and response to issues of communication in the workplace. In this context, sociolinguists are particularly well-equipped to tackle issues of power, politeness, and gender (Holmes and Stubbe, 2003; Mullany, 2003).

Cross-cultural and Intercultural

European migration toward the New World was responsible for a multilingual, multicultural workforce in North American and Australian work environments for decades, but detailed empirical studies of communication issues are still rare (Goldstein, 1997; Clyne, 2003). On the contrary, the dominant importance of trade exchange with Japan explains the steady growth of contrastive studies of (American and Australian) English and Japanese such as Miller (2000) and Yamada (1997). Intercultural business communication research in Europe has traditionally relied on Scandinavian scholarship (e.g., Neumann, 1997, Grindsted, 1997).

In one of the earliest studies of Italian and British business discourse, Bargiela-Chiappini and Harris (1997a) look at the pragma-linguistic choices of managers in the two countries but also examine the influence of cultural preferences on multiparty business meetings and on management praxis. Evidence from further empirical research suggests that a reappraisal is necessary of the role of 'culture' in business interactions (Bargiela-Chiappini and Nickerson, 2003). A possible way forward is the construct of 'interculturality,' defined as the process and the condition of cultures-in-contact, thus seeking to capture both the dynamics and the inherent or generic properties of intercultural exchanges as manifested through discourse (Bargiela-Chiappini, 2004).

Developing this argument, Poncini (2004) in her analysis of intercultural meetings convincingly argues that communicative misunderstandings are often too simplistically attributed to cultural differences and that linguistic competence (in English) alone cannot account for misunderstanding. The business context, organizational roles, and individual interactional styles are equally important in realizing and maintaining successful business relationships. The role of English as a *lingua franca* and its interplay with other languages in multicultural meetings are thus both integrated within a contextualized, dynamic, sociolinguistic, analytical framework that awaits further application on the field.

Localization and Globalization

Language needs analysis promoted the use of company surveys and questionnaires to assess which and how languages were needed, and by whom (see Coleman, 1988). Interlanguage contact continues to be a distinctive feature of European business. Indeed, Vandermeeren (1999: 276) concludes on the basis of her survey of 415 European companies that "a business interaction is rarely a monolingual event."

Research in business discourse cannot ignore recent complicating socioeconomic developments such as the internationalization of business, technological advances in telecommunications, and the rise of international management teams (Nickerson, 2000). English continues to be the privileged vehicle of management ideas as well as the medium of global trade (Rogers, 1998). However, its dominant international role has been challenged by the ascendancy of

localized forms and patterns of communication that express the tension between local preferences and global choices. For instance, in Malaysian business organizations, the normative status of 'colonial' English is giving way to the emergence of the speech forms and styles of Malaysian English (Nair-Venugopal, 2000). Similarly, communication conflicts in the Argentinean subsidiary of a European company are shown to be the product of a clash between local and global organizational cultures and identities (Gimenez, 2002).

A third perspective on intercultural communication in business is one building primarily on work in politeness theory and the notion of 'face.' From a social psychological angle, it looks at intercultural encounters through the lens of 'rapport management,' defined as "the way that language is used to construct, maintain and/or threaten social relationships" (Spencer-Oatey, 2000: 12). The usefulness of taking into account the nonverbal and the relational dimensions in the analysis of multiparty intercultural business meetings lends further support to the call for multimethod, multidisciplinary research.

Future Developments

Empirical research in work settings entails issues of access and confidentiality, which often impose severe constraints on the scope and type of fieldwork. Nonetheless, interest in the Language of Work from other than applied linguistics and sociolinguists has considerably increased. In the 1990s, organizational ethnographers and organizational discourse analysts have joined the field's varied membership (Hardy *et al.*, 2004). There seems to be no end of opportunity for crossdisciplinary, crosscultural exchange and growth; this is perhaps the most exciting prospect for a field that arguably has its roots in a single language (English) and a single setting (the classroom).

See also: Conversation Analysis; Institutional Talk; Intercultural Pragmatics and Communication; Language in the Workplace: Different Approaches; Languages for Specific Purposes; Pragmatics: Linguistic Imperialism.

Bibliography

Agar M (1985). 'Institutional discourse.' *TEXT 5(3)*, 147–168.

Bargiela-Chiappini F (2004). 'Intercultural business discourse.' In Candlin C N & Gotti M (eds.) *Intercultural aspects of specialized communication.* Bern: Peter Lang 235–248.

Bargiela-Chiappini F & Harris S (1997a). *Managing Language. The discourse of corporate meetings*. Amsterdam: Benjamins.

Bargiela-Chiappini F & Harris S (eds.) (1997b). *The languages of business: an international perspective.* Edinburgh: Edinburgh University Press.

Bargiela-Chiappini F & Nickerson C (eds.) (1999). *Writing Business: genres, media and discourses.* Harlow, UK: Longman.

Bargiela-Chiappini F & Nickerson C (2001). 'Partnership research: a response to Priscilla Rogers.' *Journal of Business Communication 38(3)*: 248–251.

Bargiela-Chiappini F & Nickerson C (eds.) (2002). *International journal of applied linguistics.* Special issue on Business Discourse.

Bargiela-Chiappini F & Nickerson C (2003). 'Intercultural business communication: a rich field of studies.' *Journal of Intercultural Studies 24(1)*: 3–16.

Boden D (1994). *The business of talk.* Cambridge: Polity Press.

Candlin S (2003). 'Issues arising when the professional workplace is the site of applied linguistic research.' Response article. *Applied Linguistics* Special Issue *24(3)*, 386–394.

Clyne M (2003). *Dynamics of language contact.* Cambridge: Cambridge University Press.

Coleman H (1988). 'Analyzing language needs in large organizations.' *English for Specific Purposes 7*, 155–169.

Drew P & Heritage J (eds.) (1992). *Talk at work. Interaction in institutional settings.* Cambridge: Cambridge University Press.

Dudley-Evans T & St John M J (1998). *Developments in English for specific purposes.* Cambridge: Cambridge University Press.

Ehlich K & Wagner J (eds.) (1995). *The discourse of international negotiations.* Berlin: Mouton de Gruyter.

Firth A (ed.) (1995). *The discourse of negotiation: studies of language in the workplace.* Oxford: Pergamon.

Gimenez J (2002). 'New media and conflicting realities in multinational corporate communication: a case study.' *International Review of Applied Linguistics in Language Teaching 40(4)*, 323–344.

Goldstein T (1997). *Two languages at work: bilingual life on the production floor.* Berlin: Mouton de Gruyter.

Grindsted A (1997). 'Joking as a strategy in Spanish and Danish negotiations.' In Bargiela-Chiappini F & Harris S (eds.) (1997b). 159–182.

Gunnarsson B-L, Linnell P & Bordberg B (eds.) (1997). *The construction of professional discourse.* London: Longman.

Hardy C, Grant D, Oswick C, Phillips N & Putman L (eds.) (2004). *Handbook of organizational discourse.* London: Sage.

Holmes J & Stubbe M (2003). *Power and politeness in the workplace: a sociolinguistic analysis of talk at work.* London: Longman.

Jablin F M & Putman L L (eds.) (2001). *The new handbook of organizational communication. Advances in theory, research and methods.* Sage: Thousand Oaks.

Miller C (2000). 'Negative assessments in Japanese–American workplace interaction.' In Spencer-Oatey H (ed.). 240–254.
Mullany L (2003). *Identity and role construction: a sociolinguistic study of gender and discourse in management.* Ph.D. dissertation, Nottingham Trent University, UK.
Nair-Venugopal S (2000). *Language choice and communication in Malaysian business.* Kuala Lumpur: Penerbit Universiti Kehangsaan Malaysia.
Neumann I (1997). 'Requests in German Norwegian business discourse: differences in directness.' In Bargiela-Chiappini F & Harris S (eds.) (1997b). 72–93.
Nickerson C (2000). *Playing the corporate language game.* Amsterdam: Rodopi.
Poncini G (2004). *Discursive strategies in multicultural business meetings.* Bern: Peter Lang.
Rogers P (1998). 'National agendas and the English divide.' *Business Communication Quarterly 61(5),* 79–127.
Sarangi S & Candlin C N (2003). 'Introduction. Trading between reflexivity and relevance: new challenges for applied linguistics. Editorial.' *Applied Linguistics* Special Issue *24(3),* 271–285.
Spencer-Oatey H (ed.) (2000). *Culturally speaking: managing rapport through talk across cultures.* London: Continuum.
Vandermeeren S (1999). 'English as a lingua franca in written corporate communication: findings from a European survey.' In Bargiela-Chiappini F & Nickerson C (eds.). 273–292.
Yamada H (1997). *Different games, different rules: why American and Japanese misunderstand each other.* Oxford: Oxford University Press.

Language Loyalty

S Bowerman, University of Cape Town, Cape Town, South Africa

'Language loyalty' refers to efforts by a speech community to maintain its own (first or home) language in the face of real or perceived threats to its status and continued use. A language may find itself threatened by one or more other languages for a host of reasons, for example: depopulation or emigration of its speech community; mass immigration of other speech communities; or coercion (martial or economic) or voluntary language shift, often for economic reasons. Because of its close association with ethnicity, loyalty to a language very often reflects loyalty to an ethnicity or culture, of which language is perceived to be the carrier (Fishman, 1997: 331f). The effects of language loyalty vary widely, from the political token use of a language to a widespread and militant language loyalty campaign. The phenomenon of language loyalty can be observed and measured in the native environment of the language in question (e.g., Ireland, where Irish has been all but supplanted by English) or in an expatriate speech community (e.g., Spanish-speaking Mexican communities in California).

The strength of a speech community's loyalty to its language largely depends on how 'conscious' the community is of its language and/or ethnicity (Fishman, 1997: 330), and, indeed, how conscious the mainstream society is of their otherness. Moreover, as Crystal (1995: 367) points out, a language community is seldom homogenous in its attitudes toward its language. Some individuals may advocate a total shift to the majority or dominant language; some may advocate functional bilingualism, in which the first language is reserved for private, home-based functions; and still others may prefer an all-out effort to promote its own language for all functions within the home base of the speech community and beyond.

The impetus for individuals and communities to resist pressures against their home language, and to continue using and developing it, springs from a variety of sources. One of these is the size and strength of the linguistic (or ethnolinguistic) community in question. The larger, more close-knit and 'ghettoized' the community is, the more pressure will be placed on individuals within it to use its language. And, indeed, there will be more opportunity to use it than in a widely scattered, more integrated speech community. Relations between the group in question and the wider society are also important; a more influential (politically and economically) group has a far better chance of being able to maintain – and promote – its language than a less influential one does.

Language Planning and Language Loyalty

While language loyalty is not necessarily a facet of language planning, the two concepts are inextricably linked. Language loyalty is very often a response to language planning, and, perhaps less often, language loyalty is expressed through a process of language planning. Status planning reflects a desire to maintain and promote the use of a threatened language in the higher functions of language use

Miller C (2000). 'Negative assessments in Japanese–American workplace interaction.' In Spencer-Oatey H (ed.). 240–254.
Mullany L (2003). *Identity and role construction: a sociolinguistic study of gender and discourse in management.* Ph.D. dissertation, Nottingham Trent University, UK.
Nair-Venugopal S (2000). *Language choice and communication in Malaysian business.* Kuala Lumpur: Penerbit Universiti Kehangsaan Malaysia.
Neumann I (1997). 'Requests in German Norwegian business discourse: differences in directness.' In Bargiela-Chiappini F & Harris S (eds.) (1997b). 72–93.
Nickerson C (2000). *Playing the corporate language game.* Amsterdam: Rodopi.
Poncini G (2004). *Discursive strategies in multicultural business meetings.* Bern: Peter Lang.
Rogers P (1998). 'National agendas and the English divide.' *Business Communication Quarterly 61(5),* 79–127.
Sarangi S & Candlin C N (2003). 'Introduction. Trading between reflexivity and relevance: new challenges for applied linguistics. Editorial.' *Applied Linguistics* Special Issue *24(3),* 271–285.
Spencer-Oatey H (ed.) (2000). *Culturally speaking: managing rapport through talk across cultures.* London: Continuum.
Vandermeeren S (1999). 'English as a lingua franca in written corporate communication: findings from a European survey.' In Bargiela-Chiappini F & Nickerson C (eds.). 273–292.
Yamada H (1997). *Different games, different rules: why American and Japanese misunderstand each other.* Oxford: Oxford University Press.

Language Loyalty

S Bowerman, University of Cape Town, Cape Town, South Africa

'Language loyalty' refers to efforts by a speech community to maintain its own (first or home) language in the face of real or perceived threats to its status and continued use. A language may find itself threatened by one or more other languages for a host of reasons, for example: depopulation or emigration of its speech community; mass immigration of other speech communities; or coercion (martial or economic) or voluntary language shift, often for economic reasons. Because of its close association with ethnicity, loyalty to a language very often reflects loyalty to an ethnicity or culture, of which language is perceived to be the carrier (Fishman, 1997: 331f). The effects of language loyalty vary widely, from the political token use of a language to a widespread and militant language loyalty campaign. The phenomenon of language loyalty can be observed and measured in the native environment of the language in question (e.g., Ireland, where Irish has been all but supplanted by English) or in an expatriate speech community (e.g., Spanish-speaking Mexican communities in California).

The strength of a speech community's loyalty to its language largely depends on how 'conscious' the community is of its language and/or ethnicity (Fishman, 1997: 330), and, indeed, how conscious the mainstream society is of their otherness. Moreover, as Crystal (1995: 367) points out, a language community is seldom homogenous in its attitudes toward its language. Some individuals may advocate a total shift to the majority or dominant language; some may advocate functional bilingualism, in which the first language is reserved for private, home-based functions; and still others may prefer an all-out effort to promote its own language for all functions within the home base of the speech community and beyond.

The impetus for individuals and communities to resist pressures against their home language, and to continue using and developing it, springs from a variety of sources. One of these is the size and strength of the linguistic (or ethnolinguistic) community in question. The larger, more close-knit and 'ghettoized' the community is, the more pressure will be placed on individuals within it to use its language. And, indeed, there will be more opportunity to use it than in a widely scattered, more integrated speech community. Relations between the group in question and the wider society are also important; a more influential (politically and economically) group has a far better chance of being able to maintain – and promote – its language than a less influential one does.

Language Planning and Language Loyalty

While language loyalty is not necessarily a facet of language planning, the two concepts are inextricably linked. Language loyalty is very often a response to language planning, and, perhaps less often, language loyalty is expressed through a process of language planning. Status planning reflects a desire to maintain and promote the use of a threatened language in the higher functions of language use

(except in cases where the use of a language is deliberately supplanted by another). For example, nine indigenous languages were elevated to official status in South Africa in 1994, alongside the two preexisting official languages, English and Afrikaans. However, English is highly dominant in this context, and there is a strong feeling among many Afrikaans speakers that Afrikaans is being deliberately supplanted by English in higher functions, particularly in the media. Corpus planning, too, is a double-edged sword: while it can be used to 'revalorize' and develop languages to make them suitable for use in higher functions (e.g., the development of the indigenous languages of South Africa for their new official status), it can also be used to strip away some part of a language's identity (Cyrillicisation of Uzbek as part of the Soviet Union's Russification programs).

Case Studies

The case studies that follow are chosen because they illustrate some very different situations in which language loyalty can be observed.

Irish: Efforts to Maintain a Language Eroded in Its Home Territory

Although Ireland and England have been involved in a contact-conflict situation for more than 800 years, English only became the language of opportunity in Ireland beginning in the mid–17th century. Previous attempts at establishing an English or Anglo-Saxon presence in Ireland had usually resulted in the Gaelicisation of the colonizers (Wardhaugh, 1988: 90). Since then, the dominance of English has been overwhelming, with the number of Irish first-language speakers dwindling to under 3% of the population. The Republic of Ireland gained independence from the United Kingdom in 1921, and in 1937 the constitution elevated Irish to the position of first official language. However, this has done little to extend the use of Irish in Ireland. A succession of revival campaigns (emphasizing loyalty to Irishness as an ethnonational identity) has met with limited success. Irish continues to fulfill token functions (first official language, offered as a subject in schools) but has very little presence in the daily life of the country as a whole. Moreover, no policies have been devised to support the 'first official language' status of Irish, and "it has taken on the iconisation of an aspiration, with no relevance for the everyday life of the country" (O'Dochartaigh, 2000). While Irish may not be the mainstream medium of communication in Ireland, its official status is becoming more and more overt: public signage is mostly bilingual, and state-sponsored radio and television stations are devoted to it.

Former Colonies in Africa: Dominated by a Minority but Powerful Language

Multilingualism is the norm in most African states. In the global context, however, this situation is starting to change under the influence of influential global languages of wider communication, such as English. People in developing countries who wish to participate on the global stage need access to these stronger languages, and this affects their loyalty to their first language. Countries emerging from colonial rule tend to opt for the former colonial language as official or 'first official' language of the country; thus, the higher functions of language use fall to, e.g., French or English. Numerous – and often not very credible – reasons are given for this. These range from the choice of a language with no tribal associations to promote national unity to the fear that the choice of an indigenous language as official language will leave the state isolated from the global scene (Webb, 1994: 194ff). A notable exception to this rule is the adoption of Swahili in Kenya (and, to a lesser extent, in other east African regions) for the higher functions.

In many developing countries with a colonial legacy, linguistic (and ethnic) diversity is perceived as a threat to the unity of the nation and to 'nation building,' and those who support ethnic languages are seen to be supporting separatism. Very little institutional support is offered to the indigenous languages, as education, media, public administration, and so forth favor only a few dominant languages (Brenzinger, 1997). The irony of this situation is that only a small segment of the population has access to the languages that serve the higher functions, while the indigenous languages remain underdeveloped as media for scientific and technical discourse. Unless substantial corpus and status planning programs are undertaken, any individual who wishes to participate in the public life of the country requires competence in the dominant languages. Thus, for many indigenous language speakers, allegiance to the wider society must come at the cost of allegiance to their home language, because the system as a whole has little allegiance to the indigenous languages of the country.

United States Immigrants: Language Maintenance Efforts in Expatriate Communities

Fishman's (1966) seminal collection of essays on language and ethnicity in the United States identifies a number of different language loyalty situations. There are a variety of immigration situations, and these affect the intensity and success of linguistic self-maintenance efforts. They can be summarized as follows:

- Immigrants who migrated to the United States for economic prosperity, and with the intention to assimilate into the culture
- Immigrants who fled a threat to their language and culture and sought to preserve it in the United States
- Immigrants who only became (self-)conscious of their ethnolinguistic identity in the United States.

There is significant overlap between the first and third categories. Those who migrated to the United States with the intention of integrating into the so-called melting-pot culture undoubtedly had few ties to their language or ethnic identity. However, having arrived in the United States without any immediate access to the mainstream language and culture, many immigrants found their ethnolinguistic identity to be an important source of stability, as a link with the old and familiar in very new surroundings. Other people and groups fled to the United States to maintain their language and culture, which were threatened in their native land – a prime example being the Ukrainian immigrants, whose home language and culture came under threat from the Russification campaign in the Soviet Union. For people in these circumstances, language maintenance is a priority (particularly for the first generation), and they actively seek out to establish a group identity. In later generations, when the ethnic consciousness is not as strong, language loyalty typically decreases somewhat and language maintenance efforts dwindle.

Conclusion: Proponents of Language Loyalty

As Fishman (1966b) points out, those who support ethnolinguistic minority maintenance in the 'supra-ethnic' age are often accused of supporting separatism and deliberately alienating themselves (and seeking to alienate others) from the mass culture. This has led to a paradoxical situation, especially in the United States. Although most historical ethnic minorities have long since abandoned their languages and most of their external cultural trappings, they still maintain a fairly high degree of ethnic self-identity (Fishman, 1966a: 29). However, there is another side to this coin, in the form of the mass immigration that continues to this day, particularly of Spanish speakers from Mexico and other parts of South America. Many among this group are functionally bilingual, and, as they edge nearer to numerical majority, are openly hostile to the U.S. *de facto* policy of English monolingualism. Pressure is mounting for the status of Spanish to be developed in affected regions, and English seems not to have the draw it used to be.

However, for relatively weak languages in the shadow of a highly dominant language, language loyalty often does not reach very far. Often, only educationalists (especially language teachers), diehard traditionalists, and linguists plug away at a language maintenance program, while the majority of actual or potential users of the language remain relatively apathetic – or express an interest in or regard for the language that is not accompanied by increased use (Fishman, 1966b: 397). A case in point here is the state of African language departments in South African universities, which have suffered a massive drop in voluntary enrollments despite the official status of nine African languages. In most cases, it seems, loyalty to a dominated language tends to be supplanted by loyalty to more materially pressing issues: economic aspirations, new identity, or even the status quo.

See also: Endangered Languages; Identity and Language; Language Maintenance and Shift; Language Planning and Policy: Models; Migration and Language; Minorities and Language; Multiculturalism and Language.

Bibliography

Brenzinger M (1997). 'Language contact and language displacement.' In Coulmas F (ed.) *The handbook of sociolinguistics.* Oxford: Blackwell. 273–284.

Crystal D (1995). *The Cambridge encyclopedia of language.* Cambridge: Cambridge University Press.

Fishman J (1966a). 'The historical and social contexts of an inquiry into language maintenance efforts.' In Fishman J (ed.) *Language loyalty in the United States.* London: Mouton. 21–33.

Fishman J (1966b). 'Language maintenance in a supra-ethnic age: summary and conclusions.' In Fishman J (ed.) *Language loyalty in the United States.* London: Mouton. 392–411.

Fishman J (1997). 'Language and ethnicity: the view from within.' In Coulmas F (ed.) *The handbook of sociolinguistics.* Oxford: Blackwell. 327–343.

O'Dochartaigh C (2000). 'Irish in Ireland.' In Price G (ed.) *Languages in Britain and Ireland.* Oxford: Blackwell. 6–36.

Wardhaugh R (1988). *Languages in competition.* Oxford: Blackwell.

Webb V (1994). 'Revalorizing the autochthonous languages of Africa.' In Pütz M (ed.) *Language contact, language conflict.* Amsterdam: John Benjamins. 181–204.

Relevant Website

http://www-rcf.usc.edu/~cmmr/ – Website of the Center for Multilingual, Multicultural Research at the University of Southern California; includes information on language policy and language rights.

Language Maintenance and Shift

M Brenzinger, University of Cologne, Cologne, Germany

On the whole, the extent of endangerment of the world's language diversity reflects the distribution of languages on our planet. The highest density of endangered languages seems to exist in the linguistically heterogeneous parts of the world. It happens that countries with very high numbers of languages are near the equator, such as Papua New Guinea, India, and Mexico, as well as the countries of central and western Africa.

An analysis of language endangerment globally shows that a crucial difference in language shift settings seems to be between contexts of globalization, nation-state contexts, and subnational contexts. In a substantial part of the world, namely Australia, the Americas, Europe, and Japan, minority languages are increasingly under threat as they are replaced by the official languages of nation-states. However, this does not hold true for the vast majority of endangered languages spoken elsewhere, for example on the African continent. With about one-third of the world's languages, this continent is among the most linguistically rich areas on our planet.

Language maintenance activities emerge in quite different contexts and facets. Some French and Afrikaners feel a pressing need for such supportive measures even though their languages seem not to be immediately threatened by extinction. In contrast, speakers of endangered minority languages may refuse to get involved in activities that aim to maintain or revitalize the use of their ethnic tongue. Among linguists, the stances toward language maintenance also vary considerably. Some seem to want to work toward a maintenance of the world's linguistic diversity, which implies a challenge to either human evolution or a creationist design. On the other extreme, some linguists do not see it as their duty to get professionally involved in language maintenance activities. In the following account, language maintenance activities will be discussed with merely a pragmatic focus, i.e., the options and roles for speech communities, linguists, and international bodies such as the United Nations Educational, Scientific, and Cultural Organization (UNESCO).

Has there ever been a *Homo sapiens sapiens* protolanguage, which, some 100 000 or 200 000 years ago, started a 12 000-year-long process of diffusing into more than 100 000 languages, which have been reduced from that number to the approximately 6000 languages spoken today? Whatever possible scenario for language evolution one chooses to believe, it seems certain that our present language inventory is the remnant of a much richer pool. Furthermore, the shrinking of linguistic diversity seems to have picked up speed in the last few thousand years. This situation is of immediate concern to many linguists and numerous ethnolinguistic communities.

Terms like 'language death,' 'language murder,' 'language suicide,' and 'killer languages' evoke pictures in which languages themselves are alive or dead. Languages are portrayed by these terms either as suffering victims or as aggressive offenders, prior to any scholarly analysis. We should keep in mind that all spoken tongues are no more and no less than variations of human language. Groups of people may foster or give up an ethnic language, the latter having happened thousands of times. Communities and individuals who give up their ethnic language do not end up in total silence, unable to communicate with anyone any longer. Such language shifters usually acquire 'modern' and more 'useful' languages. Language is one – some say the most important – asset of culture and identity. But cultures as well as identities change over time. Why then do some linguists seem to think that the status quo of language usage could and should be frozen?

While scholars would do well to refrain from sentimental statements on languages of other people, they may very well respect and assist in spreading the emotional attachment of the speakers themselves to their language. Minorities and their offspring very often experience the loss of their heritage languages as traumatic. For example, as First Nations people in Canada phrase it, without an ancestral language, people are no longer able to relate to their environment or to their spiritual world. Khwe in Namibia, Bretons in France, Ainu in Japan, Frisians in the Netherlands, Mohawks in Quebec, and many other ethnolinguistic minorities feel that way and want to maintain or regain their languages. With such communities, the question arises, what can be done and by whom?

The Intangible Cultural Heritage Unit of UNESCO has been working for more than 20 years with members of endangered language communities and linguists toward establishing an Endangered Language Section. These initiatives have resulted in the formulation of the documents 'Language Vitality and Endangerment' as well as 'Recommendations to the Director General' labeled 'Action Plans' in March 2003. I will highlight some of the key issues discussed in these documents.

It seems to me that there are basic differences in the types of language endangerment in different language contact settings. While we are quite familiar with the language endangerment situation in nation-state contexts, as found for example in the Americas, Australia, and Europe, we still know little about the other parts of the world. The greatest linguistic diversity still exists in Papua New Guinea and in Africa south of the Sahara. The threat to these languages differs from that in the previously mentioned parts of the world.

Ethnolinguistic minorities from nation-state contexts, such as the First Nations people in Canada or the Amazigh in Morocco, may benefit from examining the living experience and common practice of multilingualism in most sub-Saharan African communities. The latter, in turn, must learn about the experiences of minorities in nation-states because this will be their challenge in the near future. UNESCO is playing an important role by supporting African governments in establishing language policies that might help to allow for cultural and linguistic diversity within their nation-states.

The role of UNESCO in that process is threefold. UNESCO addresses governments on their educational and cultural policies within the frame of its 'Universal Declaration on Cultural Diversity' of 2001. But UNESCO in addition requests the active involvement of linguists and speech communities to safeguard the still-existing linguistic diversity. Noriko Aikawa, director of the Intangible Heritage Section of UNESCO, stated in 2001: "UNESCO's future strategy on languages is to strengthen co-ordination between language research, policy and education programs. The results of research should be immediately incorporated into appropriate language policies, educational planning and pedagogical tools." (Aikawa, 2004: 35).

Languages Are Threatened in Different Settings

The scenarios in which the world's language diversity is threatened range from subnational settings to national contexts and finally to environments that are formed by processes associated with globalization. It is crucial, indeed, for the understanding of various aspects of language endangerment, to consider the specific contexts in which language shifts occur. In the following discussion, three different settings of language contact situations will be distinguished: the global context, the context of nation-states, and the subnational context.

Global Context

Looking at the dynamics by which languages gain speakers in the context of globalization, we find that only a few languages benefit from global merging, and these languages at present seem to be mainly English and Spanish. The international exchange of knowledge and world trade are conducted to an increasing degree in only a few world languages, and some scholars seem to expect that, given this tendency, a world culture, based on one common language, will finally emerge from these developments.

But up to now, we can state that languages have not been replaced in the context of globalization. The so-called world languages – increasingly English only – are acquired as additional and not as first languages. For example, the integration of the new European Union member states accelerates the spread of English as the European lingua franca. In the course of this process, French will further lose second language speakers, but at the same time, none of the European national languages, especially not French, will lose first language speakers to English ('After Babel,' 2004).

Nation-State Context

In the context of nation-states, governments obviously rely on the languages in which they run the countries. National majority languages are widely established and used in administration, politics, science, and education, as well as in media and literature. Such national languages may not be confined to single political-administrative territories; the same language may be in such a dominant position in several nation-states. National languages may also not cover an entire nation-state, as we see in Quebec. However, common to all these national languages is the criterion that they are instrumentalized by governments, and with that they receive official support and recognition within nation-states.

In the same context of nation-states, we often find minority languages, which are threatened and marginalized by the dominance of established national languages. Languages of this latter type are, for example, Japanese, Finnish, German (German, Standard), and Dutch. In their areas of domination they endanger minority languages such as Ainu, Saami, Sorbic, and Frisian, respectively. At the same time, various forms of English are threatening and replacing minority languages in quite a number of nation-states, such as the United States, Canada, the United Kingdom, Australia, and New Zealand. Within north African states, national Arabic varieties

are expanding – not only as a second language but also as a mother tongue – and replacing the ancestral languages of Amazigh communities.

Kiswahili (Swahili), in the Republic of Tanzania, and Setswana (Tswana), in the Republic of Botswana, are among the few African languages that are established as media of nationwide communication. Only these two sub-Saharan African languages have been described as threats to other languages in the context of the nation-state. Kiswahili threatens more than 130 other Tanzanian languages, while Setswana does the same to about 30 languages spoken in Botswana. Other African languages with a nationwide distribution, such as Kinyaruanda (Rwanda) and Kirundi (Rundi), do not threaten minority languages at all, because no minority languages exist within the national boundaries.

Subnational Contexts

Minority languages in many parts of the world, however, are threatened in subnational contexts. In sub-Saharan Africa, as just mentioned, very few African languages are established nationwide. On the national level, African governments predominantly use the imported languages of the former colonial powers, and these are the prevailing languages in national administration, in secondary and higher education, in modern literature, etc. In most African countries, the knowledge of these imported languages is still – after 40 years of independence – confined to the educated elite in the urban centers. One indication that imported, formerly colonial languages have not developed into the languages of the masses is the fact that not a single African language has so far been replaced by French and English. The only exception on the African continent seems to be the language of the Guanche on the Canary Islands, who lost their Amazigh language to Spanish after their final defeat in 1496. If we consider Afrikaans as a European language, then of course quite a number of Khoisan languages have been replaced by that Indo-European language, which is spoken exclusively on the African continent.

For that reason, language displacement in sub-Saharan Africa occurs, in by far the most cases, in subnational contexts, i.e., minority languages are generally not replaced by established national majority languages. Language shift takes place in local contact situations and for quite different reasons. For example, speakers of a minority language may shift from their language to another due to the adoption of a different religious faith. Former hunter-gatherer communities may assimilate to pastoral societies not only in the mode of production but also by taking the language of the pastoralist. Thus, African minority languages are threatened and replaced by other African languages, languages which themselves are very often minor languages within the nation-state.

Ethnolinguistic minorities in sub-Saharan Africa generally live in complex language contact situations in which they are marginalized in subnational contexts and for various reasons. The processes that may lead to the extinction of their languages are very different from those that affect minority languages in the contexts of nation-states. In the latter context, minority languages are generally threatened because they are dominated by national majority languages. In sub-Saharan Africa, in contrast, most languages of ethnolinguistic minorities have survived until today precisely because their speakers have been and are being marginalized and neglected, in other words, excluded from national developments. Of course, this should not lead us to be cynical and to conclude that it may be better to let the people suffer in order to keep their languages alive, so to speak. With continued efforts and good fortune, the sub-Saharan African nations, too, will progress further, and nation-states there will become the dominant frames of reference for minorities and their cultures. Then, policies that respect and appreciate cultural and linguistic diversity as an asset rather than an obstacle to national unity may help to foster the use of heritage languages among minority communities. This seems to be one of the most prominent fields for further activities for UNESCO and also for linguists.

Which Languages Are Endangered?

The evaluation of the state of vitality of any language is a challenging task, mainly because speech communities are complex and language use patterns within these communities are difficult to explore. In addition, poor infrastructure and rigid political conditions may not even allow for determining the number of actual speakers of a language.

In 2001, an Ad-hoc Expert Group on Endangered Languages was asked by the Intangible Cultural Heritage Unit of UNESCO to prepare a draft overview of the threat to the world's language diversity. From March 10 to 12, 2003, an International Expert Meeting on Endangered Languages was held at the UNESCO headquarters in Paris, and linguists and language planners, representatives of nongovernmental organizations, and members of endangered-language speech communities discussed the draft. Two documents, namely, 'Language Vitality and Endangerment' and recommendations to the director general (labeled 'Action Plans'), were formulated in March 2003.

Table 1 Factors for assessing endangered languages[a]

Degree of endangerment	
1. Intergenerational language transmission	
2. Absolute number of speakers	
3. Proportion of speakers within the total population	
4. Loss of existing language domains	
5. Response to new domains and media	
6. Material for language education and literacy	
7. Governmental and institutional language attitudes and policies, including official status and use	
8. Speakers' attitudes towards their own language	
Urgency of documentation	
9. Amount and quality of documentation	

[a]UNESCO, 2003.

In the first document, nine major factors are identified to assess the language situation of endangered languages (see Table 1).

Factors 1 through 6 focus on the assessment of a language's vitality and its state of endangerment. The single most important factor is the first one, which asks for the extent of language acquisition among the children within a community. It is obvious that a language without any young speakers is seriously threatened with extinction.

Factors 1 through 5 are meant to capture the dynamics of the processes of a given language shift situation. The proportion of speakers within a community (factor 3) reveals an important aspect of language vitality: is the minority language an essential indicator for being regarded as a member of the community or not? The Arggobba in Ethiopia, for example, strongly insist on and declare their ethnic identity and origin even though their language has disappeared and it is not known whether any speakers of Arggobba still exist among the more than 60 000 Arggobba.

Loss of domains (factor 4) for an ethnic tongue may be triggered by, for example, the introduction of formal education or by new job opportunities for members of minority groups. A change in religious affiliation can also result in a shift to another mother tongue, a language that is associated with the new religion (factor 5). For example, Hausa and Dyula (Jula) spread in west Africa along with Islam. An outstanding example of the correlation between language and religious affiliation is the community of the Jeri leatherworkers, who live among the Sienare Senofo in the northern part of the Ivory Coast. The members of the Jeri community abandon their own language and adopt two different languages as new mother tongues according to their religious affiliation: "Sienare with the non-Muslim, and Manding with the Muslim Jeri" (Kastenholz, 1998: 259).

Factor 6 relates to the stage of development of a given language ('Ausbau'). Is there a community orthography? Have they agreed on a common standard form for writing the language? Do teaching and learning materials for the language exist? Is literature, such as newsletters, stories, religious texts, and so forth, published in that language?

Factor 8 assesses the speakers' attitudes toward their ethnic language, and factor 7 deals with the government's policies toward that language. And finally, factor 9 attempts to evaluate the urgency for documentation by asking for the quantity and quality of existing analyzed language data.

For minority languages in national contexts, such as in Canada, New Zealand, and Morocco, it is obvious which of the languages have to be regarded as being endangered, i.e., the First Nations languages, Maori, and Amazigh, respectively. In other contexts, all these nine factors should be taken together in surveying the overall sociolinguistic situation of a language with respect to its degree of endangerment.

What Is Actually Lost When Languages Vanish?

The preamble to UNESCO's 'Language Vitality and Endangerment' document includes the following statement:

> The extinction of each language results in the irrecoverable loss of unique cultural, historical, and ecological knowledge. Each language is a unique expression of the human experience of the world. Thus, the knowledge of any single language may be the key to answering fundamental questions of the future. Every time a language dies, we have less evidence for understanding patterns in the structure and function of human language, human prehistory, and the maintenance of the world's diverse ecosystems. Above all, speakers of these languages may experience the loss of their language as a loss of their original ethnic and cultural identity. (UNESCO, 2003: 2, with permission.)

Obviously, linguists need language data from as many languages as possible in order to reconstruct language history, set up typologies on aspects of structure, or understand concepts underlying languages. But above that, do endangered languages offer anything special to linguists?

In his article 'The endangered languages of Ethiopia: what's at stake for the linguists?' Hayward (1998) provides several examples of unusual structural phenomena that exist only in endangered Ethiopian languages. For example, an established typology of number systems had to be fundamentally revised because of linguistic forms found in

Bayso (Baiso), an endangered language spoken in the southern part of Ethiopia.

The lexicon and categories of each language are based on and reflect a certain conceptualization of the world. The human population shares more and more perceptions of the world channeled through a rapidly spreading global culture. With that, concepts diverging from the global norms are rapidly fading and appear to be the most vulnerable assets of languages. The numeral systems, spatial orientation, and taxonomies are only some of the areas that are being homogenized in this process (Brenzinger, 2003).

As a final example, even the sounds of languages may no longer be heard, which is unquestionably an essential loss, maybe not only for linguists. The bilabial click, also referred to as the 'kiss click,' had been reported by the late Oswin Köhler (personal communication) to be extinct with the death of the last speaker of a Southern Khoesan language in the early 1970s. At the same time, Anthony Traill collected language data from !Xóõ, "the last of the Southern Bushmen languages that were once spoken throughout southern Africa" (Traill, 1985: 6). With these approximately 1000 speakers of variants of !Xóõ, a unique sound of human language, the kiss click, has survived until today.

In turning to the speakers themselves, we may ask: do they suffer from any loss when they abandon their heritage language?

Lee Cronk, an American social anthropologist, has studied the change in ethnicity and culture of a former hunter-gatherer people in Kenya. This community was formerly known as Yaaku, then Mukogodo, and today they belong to the Maasai society. The community lost the former Cushitic language and adopted the Maa language. The Mukogodo took over the pastoral mode of production and along with that the culture of the Maasai.

> Mukogodo [born around the 1930s] and before were born in rock shelters, grew up speaking Yaaku, and ate honey and wild game. Younger people were born in small houses, grew up speaking Maa, and drink milk and eat meat from goat, sheep, and cattle. In the past few decades, Mukogodo have also begun attending schools, getting jobs in and outside the Mukogodo area, and even traveling to Ulaya [Europe] itself. Change has been the norm for most humanity for the past century or more, and nowhere has it been faster or more dramatic than among the Mukogodo. (Cronk, 2004: 2, with permission.)

Maasai culture has spread over a wide area of east Africa, and several hunter-gatherer communities have lost their distinct identity and language in the "homogenizing force of Maasai society and culture" (Cronk, 2004: 143). Cronk captured an experience I had while working with a Yaaku elder in 1990:

> One day while conducting an interview with an elderly Mukogodo man about the old language [Yaaku], Brenzinger played back some Yaaku phrases that the man had spoken earlier into a tape recorder. At first, the man was delighted to hear the old language, even if it was coming from a box rather than from a person. Not really understanding the principle of the tape recorder, he tried engaging it in conversation, giving the customary cheery 'Eiuwuo!' response to the recorded greeting 'Aichee!' But soon it became obvious that the tape recorder did not really know how to carry on a conversation in Yaaku. The man first became frustrated, then very angry, and finally he began to weep. (Cronk, 2004: 83, with permission.)

In 1990, only a few elders remembered fragments of the old Yaaku language, and the emotional affection for the Yaaku language, as expressed in the above story, was not shared by other community members. The young members of the community were not interested in the language of their grandparents, whom they thought of as 'primitive' cave people, even though some of them were still alive. The community had adopted the Maasai lifestyle and values in sharing the high appreciation for cattle and Maasai customs such as female circumcision, the age-set system (in which groups of men go through initiation and other stages of life together), and marriage ceremonies.

Although the Mukogodo-Maasai decided to abandon their former language and culture, many other indigenous peoples insist that linguistic (and ethnic) identity is essential and enables them to respond to the everyday challenges in their lives. The deep-rooted emotional affection of many speakers for their language has been expressed very often. The following statement is by Christine Johnson, a Tohono O'odham elder, for the American Indian Language Development Institute in June 2002.

> I speak my favourite language
> because
> that's who I am.
> We teach our children our favourite language,
> because
> we want them to know who they are. (UNESCO, 2003: 1)

There are members of various globally dominant cultures who think that the worldwide spread of one language, of course their own (together with their culture, economy, and religion), would be for the benefit of all humankind. Obviously, not everyone agrees on that. However, one also has to respect the

will of communities that decide to abandon their language and culture, as described with the former Yaaku, today's Mukogodo-Maasai.

What Could and Should Be Done?

What is the first priority when languages are threatened by extinction? Among scholars, there is hardly any argument: documentation, i.e., research on and collection of data from endangered languages, is the fundamental task for linguists working in this field.

Within ethnolinguistic communities discussions on the future of their ancestral languages are far more complex, and quite diverse opinions are expressed. Those speaking endangered languages often consider their own language as being backward and not functional for themselves and future generations. Other communities, however, experience threats to their languages as a crisis and commit themselves to language (re)vitalization activities, establishing environments, such as kindergartens, in which their languages are exclusively spoken in order to stabilize their mother languages among the young generation. An increasing number of ethnolinguistic minorities want more. Firstly, they demand control over the terms and conditions that govern the conduct of research; secondly, they claim rights on research outcomes and also want to have a say in how research results should be used and disseminated.

There is a minimal consensus among linguists on the overall importance of the preservation of language data of endangered languages. However, reactions to demands for assistance by the speech communities with which linguists work may vary widely. Such requests relate mainly to four areas, which may be regarded as essential for safeguarding endangered languages.

Language Work

Many scholars working with ethnolinguistic minorities are willing to help develop practical orthographies that can be used by community members themselves. And some scholars assist communities in the production of reading, learning, and teaching materials.

Capacity Building

Capacity building in this context may refer to the training of local language workers in reading and writing their own languages, the production of reading materials, etc. More privileged settings further allow for the training of language teachers. The formation of local, academic research centers in which speakers of endangered languages are trained to study and document their own languages is possible with quite a few ethnolinguistic minorities. And an exceptional case for the African continent is the Royal Institute for the Amazigh Culture, in which professors and researchers have worked since October 17, 2001, in a national research institute at the University of Rabat to document and develop their own threatened mother tongue, Amazigh.

Language Policy

Language policy is a highly sensitive issue in well-established democracies and even more so in countries that are still struggling to find their way to good governance. Very few linguists are actively involved in the formulation of national language policies in countries of the first kind, and none I know of in those of the second.

In the educational sector, quite a number of linguists are engaged in implementing mother tongue education programs to safeguard ancestral languages. Mother tongue education has become more popular in most parts of the world over the past 15 years, and UNESCO has been instrumental in this development through its policy statements and related activities since 1953.

Looking at endangered languages, however, we find that in many Asian and African countries, so-called mother tongue education does not refer to the ancestral languages of ethnolinguistic minorities but to the use of local, provincial, and national dominant languages as the media of instruction. Less than 10% of the approximately 2000 African languages are currently employed as media of instruction in the educational sector, with not even a single endangered language among them. Mother tongue education in many cases further cements the position of languages that spread at the expense of endangered languages. As linguists, we are obliged to support any attempt to use African languages in formal education, but with that we may involuntarily help to threaten the languages of ethnolinguistic minorities, which are not included among the media of mother tongue education. There seems to be no way out of this dilemma.

Living Conditions

Dealing with the living conditions of communities is commonly not considered to be within the assignment of linguists. Nevertheless, linguists could help to overcome problems caused by economic poverty and lack of education. For example, national HIV/AIDS awareness programs do not generally consider the often-illiterate ethnolinguistic minorities. Materials

produced for these marginalized communities require input from linguists because the concepts and contents need to be conveyed in a culturally meaningful way. Only then can these communities understand the challenge of HIV/AIDS and react to it appropriately.

Roles for Linguists and Speech Communities

Maintaining language diversity demands the involvement of linguists, language planners, and policy makers. Akira Yamamoto has been rightfully demanding for many years that 'research in endangered language communities must be reciprocal and collaborative' because only in working together with the communities are linguists able to contribute to the safeguarding of endangered languages.

Some scholars estimate that in the last 500 years at least half of the languages formerly spoken throughout the world have disappeared. They suggest that only 10% of the present day languages are safe, i.e., not threatened with extinction. To safeguard the vast number of endangered languages requires actions in quite a number of different areas, far more than the few mentioned above. The specific activities that can and should be carried out in a given research setting are determined by several factors. Each research context has its specific limitations and prospects, and these always depend heavily on communities' attitudes toward their own language and their perception of the research carried out on it. And yet, the professional and social competence of a researcher is crucial in determining what is actually possible in the collaboration between the scholar and the speech community.

Having a linguist who is a member of the community does not always keep other community members from doubting linguistic research being carried out on an endangered language. Ofelia Zepeda, a trained linguist who wrote her Ph.D. thesis on her mother tongue, describes the most critical challenges to her linguistic work, which came from her own people, the O'odham. The objections to her work came from two different angles. The few O'odham people who had received some linguistic training some years before her, instead of supporting her, created obstacles to her work.

> Interestingly it was this small group of skilled O'odham language teachers who looked at my publication critically and in their own way offered criticism. However, their criticism was often not constructive or friendly. While still other O'odham people, taking the conservative position, let me know that it was wrong to write the language and to publish it for wide dissemination. (Zepeda, 2004: 7)

Dr Zepeda published *A Papago grammar* in 1982, but the O'odham people have only now, after more than 20 years, started to acknowledge her academic and community work on her mother tongue.

Proficiency in nationally and internationally dominant languages will gain importance throughout the world and, for that reason, will continue to spread. This development does not necessarily require the sacrifice of other languages, i.e., mother tongues of ethnolinguistic minorities, because most societies have always been multilingual. However, speakers might decide to abandon a low-prestige ethnic tongue for the benefit of social mobility and career opportunities. Languages are either maintained or abandoned by their speakers. In these situations, ancestral languages can survive in the long run only if meaningful roles for them can be established in the lives of the community members. Ultimately, in order to maintain and perpetuate the world's language diversity, these speakers have to find good reasons for keeping their ancestral language alive in natural everyday communication with their offspring.

See also: Endangered Languages; Environment and Language; Ethnolinguistic Vitality; Identity and Language; Language Planning and Policy: Models; Migration and Language Planning; Mother Tongue Education: Nonstandard Language; Reversing Language Shift; Tohono O'odham.

Bibliography

'After Babel, a new common tongue. It turns out to be English.' (2004). *Economist* August 7, 23–24.

Aikawa N (2004). 'UNESCO's programme on languages.' In Sakiyama O, Endo F, Watanabe H *et al.* (eds.) *Lectures on endangered languages: 4 – from Kyoto conferences 2001*. Osaka: ELPR. 29–35.

Bradley D & Bradley M (eds.) (2002). *Language endangerment and language maintenance*. London: Routledge Curzon.

Brenzinger M (ed.) (1992). *Language death: factual and theoretical explorations with special reference to east Africa*. Berlin: Mouton de Gruyter.

Brenzinger M (1992). 'Lexical retention in language shift: Yaaku/Mukogodo-Maasai and Elmolo/Elmolo-Sanmburu.' In Brenzinger M (ed.) *Language death: factual and theoretical explorations with special reference to east Africa*. Berlin: Mouton de Gruyter. 213–254.

Brenzinger M (1997). 'Language contact and language displacement.' In Coulmas F (ed.) *The handbook of sociolinguistics*. Oxford: Blackwell. 273–284.

Brenzinger M (ed.) (1998). *Endangered languages in Africa*. Cologne: Rüdiger Köppe.

Brenzinger M (1998). 'Various ways of dying and different kinds of deaths: scholarly approaches to language

endangerment on the African continent.' In Matsumura K (ed.) *Studies in endangered languages.* Tokyo: Hituzi Syobo. 85–100.

Brenzinger M (2003). 'Language and conceptual diversity under threat: language endangerment on the African continent.' In Vielberth J & Drexler G (eds.) *Linguistic cultural identity and international communication. Maintaining language diversity in the face of globalization.* Saarbrücken: AQ-Verlag. 59–77.

Cronk L (2004). *From Mukogodo to Maasai: ethnicity and cultural change in Kenya.* Boulder: Westview.

Crystal D (2000). *Language death.* Cambridge: Cambridge University Press.

Dorien N C (ed.) (1989). *Investigating obsolescence: studies in language contraction and death.* Cambridge: Cambridge University Press.

Dwyer A, Brenzinger M & Yamamoto A (2003). 'Safeguarding of endangered languages: report on the project of the Intangible Cultural Heritage Section of UNESCO.' *ELFN (Endangered Language Fund Newsletter) 7(1),* 1–4.

Fishman J A (ed.) (1991). *Reversing language shift: theoretical and empirical foundations of assistance to threatened languages.* Clevedon: Multilingual Matters.

Fishman J A (ed.) (2000). *Can threatened languages be saved? reversing language shift, revisited: a 21st century approach.* Clevedon: Multilingual Matters.

Grenoble L A & Whaley L J (eds.) (1998). *Endangered languages.* Cambridge: Cambridge University Press.

Hayward R J (1998). 'The endangered languages of Ethiopia: what's at stake for the linguists?' In Brenzinger M (ed.) *Endangered languages in Africa.* Cologne: Rüdiger Köppe. 17–38.

Hilton L & Hale K (eds.) (2001). *The green book of language revitalization in practice.* San Diego: Academic Press.

Janse M & Tol S (eds.) (2003). *Language death and language maintenance: theoretical, practical and descriptive approaches.* Amsterdam: John Benjamins.

Kastenholz R (1998). 'Language shift and language death among Mande blacksmiths and leatherworkers in the diaspora.' In Brenzinger M (ed.) *Endangered languages in Africa.* Cologne: Rüdiger Köppe. 253–266.

Krauss M (1992). 'The world's languages in crisis.' *Language 68(1),* 4–10.

Krauss M (1998). 'The scope of the language endangerment crisis and recent response to it.' In Matsumura K (ed.) *Studies in endangered languages.* Tokyo: Hituzi Syobo. 101–113.

Nettle D & Romaine S (2000). *Vanishing voices: the extinction of the world's languages.* Oxford: Oxford University Press.

Newman P (2003). 'The endangered languages issue as a hopeless cause.' In Janse M & Tol S (eds.) *Language death and language maintenance: theoretical, practical and descriptive approaches.* Amsterdam: John Benjamins. 1–13.

Matsumura K (ed.) (1998). *Studies in endangered languages.* Tokyo: Hituzi Syobo.

Robins R H & Uhenbeck E M (eds.) (1991). *Endangered languages.* Oxford: Berg.

Sakiyama O, Endo F, Watanabe H *et al.* (eds.) (2004). *Lectures on endangered languages: 4 – from Kyoto conferences 2001.* Osaka: ELPR.

Traill A (1985). *Phonetic and phonologocal studies of the !Xóõ, bushman* [Quellen zur Khoisan-Forschung, 1]. Cologne: Rüdiger Köppe.

UNESCO Ad Hoc Expert Group on Endangered Languages (2003). 'Language Vitality and Endangerment. Document submitted to the International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages.' Paris, 10–12 March.

Wurm S A (ed.) (2001). *Atlas of the world's languages in danger of disappearing* (2nd edn.). Paris: UNESCO Publishing.

Zepeda O (1982). *A Papago grammar.* Tucson: University of Arizona Press.

Zepeda O (2004). 'Linguistic research at home: making it our own, for our own.' In Sakiyama O, Endo F, Watanabe H *et al.* (eds.) *Lectures on endangered languages: 4 – from Kyoto conferences 2001.* Osaka: ELPR. 3–11.

Language of Legal Texts

P Tiersma, Loyola Law School, Los Angles, CA, USA

Introduction

A legal text differs in a number of ways from ordinary spoken language. This is especially true of authoritative legal texts: those that create, modify, or terminate a legal state of affairs. Such texts are what J. L. Austin might have called written performatives. Lawyers often refer to them as operative or dispositive. Although the focus of this article is on Anglo-American law, the general nature of legal texts is remarkably similar across legal systems, although the details of their structure and use can vary considerably.

The Development of the Text

Legal transactions must originally have been performed by means of speech, often accompanied by rituals of various types. With the development of literacy, it became possible to make a record of such transactions in writing or even to transact them in

endangerment on the African continent.' In Matsumura K (ed.) *Studies in endangered languages.* Tokyo: Hituzi Syobo. 85–100.
Brenzinger M (2003). 'Language and conceptual diversity under threat: language endangerment on the African continent.' In Vielberth J & Drexler G (eds.) *Linguistic cultural identity and international communication. Maintaining language diversity in the face of globalization.* Saarbrücken: AQ-Verlag. 59–77.
Cronk L (2004). *From Mukogodo to Maasai: ethnicity and cultural change in Kenya.* Boulder: Westview.
Crystal D (2000). *Language death.* Cambridge: Cambridge University Press.
Dorien N C (ed.) (1989). *Investigating obsolescence: studies in language contraction and death.* Cambridge: Cambridge University Press.
Dwyer A, Brenzinger M & Yamamoto A (2003). 'Safeguarding of endangered languages: report on the project of the Intangible Cultural Heritage Section of UNESCO.' *ELFN (Endangered Language Fund Newsletter) 7(1),* 1–4.
Fishman J A (ed.) (1991). *Reversing language shift: theoretical and empirical foundations of assistance to threatened languages.* Clevedon: Multilingual Matters.
Fishman J A (ed.) (2000). *Can threatened languages be saved? reversing language shift, revisited: a 21st century approach.* Clevedon: Multilingual Matters.
Grenoble L A & Whaley L J (eds.) (1998). *Endangered languages.* Cambridge: Cambridge University Press.
Hayward R J (1998). 'The endangered languages of Ethiopia: what's at stake for the linguists?' In Brenzinger M (ed.) *Endangered languages in Africa.* Cologne: Rüdiger Köppe. 17–38.
Hilton L & Hale K (eds.) (2001). *The green book of language revitalization in practice.* San Diego: Academic Press.
Janse M & Tol S (eds.) (2003). *Language death and language maintenance: theoretical, practical and descriptive approaches.* Amsterdam: John Benjamins.
Kastenholz R (1998). 'Language shift and language death among Mande blacksmiths and leatherworkers in the diaspora.' In Brenzinger M (ed.) *Endangered languages in Africa.* Cologne: Rüdiger Köppe. 253–266.
Krauss M (1992). 'The world's languages in crisis.' *Language 68(1),* 4–10.
Krauss M (1998). 'The scope of the language endangerment crisis and recent response to it.' In Matsumura K (ed.) *Studies in endangered languages.* Tokyo: Hituzi Syobo. 101–113.
Nettle D & Romaine S (2000). *Vanishing voices: the extinction of the world's languages.* Oxford: Oxford University Press.
Newman P (2003). 'The endangered languages issue as a hopeless cause.' In Janse M & Tol S (eds.) *Language death and language maintenance: theoretical, practical and descriptive approaches.* Amsterdam: John Benjamins. 1–13.
Matsumura K (ed.) (1998). *Studies in endangered languages.* Tokyo: Hituzi Syobo.
Robins R H & Uhenbeck E M (eds.) (1991). *Endangered languages.* Oxford: Berg.
Sakiyama O, Endo F, Watanabe H *et al.* (eds.) (2004). *Lectures on endangered languages: 4 – from Kyoto conferences 2001.* Osaka: ELPR.
Traill A (1985). *Phonetic and phonologocal studies of the !Xóõ, bushman* [Quellen zur Khoisan-Forschung, 1]. Cologne: Rüdiger Köppe.
UNESCO Ad Hoc Expert Group on Endangered Languages (2003). 'Language Vitality and Endangerment. Document submitted to the International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages.' Paris, 10–12 March.
Wurm S A (ed.) (2001). *Atlas of the world's languages in danger of disappearing* (2nd edn.). Paris: UNESCO Publishing.
Zepeda O (1982). *A Papago grammar.* Tucson: University of Arizona Press.
Zepeda O (2004). 'Linguistic research at home: making it our own, for our own.' In Sakiyama O, Endo F, Watanabe H *et al.* (eds.) *Lectures on endangered languages: 4 – from Kyoto conferences 2001.* Osaka: ELPR. 3–11.

Language of Legal Texts

P Tiersma, Loyola Law School, Los Angles, CA, USA

Introduction

A legal text differs in a number of ways from ordinary spoken language. This is especially true of authoritative legal texts: those that create, modify, or terminate a legal state of affairs. Such texts are what J. L. Austin might have called written performatives. Lawyers often refer to them as operative or dispositive. Although the focus of this article is on Anglo-American law, the general nature of legal texts is remarkably similar across legal systems, although the details of their structure and use can vary considerably.

The Development of the Text

Legal transactions must originally have been performed by means of speech, often accompanied by rituals of various types. With the development of literacy, it became possible to make a record of such transactions in writing or even to transact them in

written form. Indeed, the writing of contracts, wills, and codes of laws was an early function of literacy in many different parts of the world.

The Evolution of Private Legal Texts

In England, the birthplace of the common law that is used throughout the English-speaking world, an early type of legal text was the will, as well as grants of land and other types of property. Throughout much of the Anglo-Saxon period, wills and grants of land were made orally, usually with witnesses present, and accompanied by some kind of ritual, such as the symbolic handing over of a clod of dirt to the purchaser or recipient. After England's conversion to Christianity, a class of literate clergymen emerged. Before long, clerics began to make written records of these wills and property transfers as memorials of the transaction, especially when the gift was made to a religious institution that had an interest in preserving evidence of the transaction in case of a future dispute.

These early texts are quite clearly evidentiary rather than performative or dispositive. The transaction was performed orally in front of witnesses. If a written document was made, it was merely evidence of that oral ceremony, not an operative or dispositive legal document in the modern sense. This is evident from the fact that, for the most part, the scribes used verbs that were in the third person and the past tense rather than the first-person present tense verbs that are commonly used performatively, as the following introductory phrase illustrates:

> HER is geswutelod an ðis gewrite hu Ælfheah ealdorman his cwidæ gecwæðan hæfð . . .
> 'Here in this document it is declared how the ealdorman Ælfheah has declared his will. . .'

In the late Anglo-Saxon period, however, there are indications of a new trend in which legal texts were no longer merely evidence of land transfers or wills. Over the centuries, it became the writing and signing of the document (or placing a seal on it) that constituted the legally operative or dispositive act. The legal transaction was now performed by means of the text. Eventually, a written text came to be considered an essential ingredient of the transaction. The final critical development in the testamentary text was the parol evidence rule, which in the case of wills prohibits the introduction of evidence of oral discussions regarding the content of the text, subject to some narrow exceptions. Thus, what the maker of the will (usually called the testator) might have said to the lawyer who drafted the will, or what he might have said to anyone else about the contents of the will, became largely irrelevant. The evolution from oral to written transaction was now complete. What mattered now was solely the written text. That text was not just the best evidence of the legal transaction. It was not just the only evidence. At this stage, the text had **become** the transaction. A will today is a specific type of legal text that is taken as the definitive expression of the testamentary intentions of its maker.

The development of English wills is a good illustration of how legal texts are introduced into a culture. In the beginning, such texts are almost always evidentiary, serving as nothing more than records of an oral transaction. Yet the text can also become authoritative, as have wills in the past few centuries. In that case, the transaction in question can only be performed by properly enacting written text.

Not all private legal transactions require authoritative written texts. Contracts are a good example. In fact, in the law of contracts we can see all three major stages in the progression from the exclusively oral to the authoritative written text. Even today, there is no requirement that a contract be in writing. A contract can be completely oral, just as in Anglo-Saxon England. It is also possible to make an oral contract that is memorialized by a written record that contains some or all of its terms. If there is a subsequent dispute about the agreement, what the parties said to one another will be crucial; the writing is simply evidence of the transaction.

Finally, the parties to a contract might decide that they want to create an authoritative text of their agreement. Lawyers call this an integrated agreement. If the parties decide to reduce the entire agreement to writing, any terms that are not included in that writing become legally irrelevant. When a court needs to decide what the agreement was between the parties, it is supposed to look only at the text itself. In other words, a text that is considered a fully integrated agreement is not just evidence of the contract, but has become the contract itself.

Public Texts

The evolution of public legal texts, including statutes and judicial opinions (called 'judgments' in England), was similar to that of private texts. Directly after the arrival of Christian missionaries in England, codes or compilations of laws began to appear, usually in the name of the king who had them written down. For the most part, the early codes did not make new law but, rather, created a record of law that already existed in people's minds or memories. Such texts, in other words, were mostly evidentiary.

Legislation, in the sense of enacting innovative rules of behavior, was rare in England until the end of the 13th century. Even then, the texts of such early

legislation were not authoritative in the modern sense; the drafting of the text of early statutes was typically done by clerks or judges after Parliament had given its assent. The text of legislation at this stage was merely evidence of a decision that Parliament had made. As a result, statutes were quite freely interpreted. Without authoritative and accurate texts, it is difficult and dangerous to focus too much on the words of the statute.

But toward the end of the 15th century, legislation in England underwent an important shift. Parliament began to enact text that had been written out beforehand. The text of statutes was no longer simply evidence of the law but began to constitute the law itself. The verbs no longer referred to what the king "hath ordained" or what Parliament "have established" but began to speak in the present tense of what the king and Parliament are enacting by means of the statute.

Statutes also illustrate another important distinction between evidentiary and authoritative texts. Earlier statutes were essentially a type of transcript or minutes written by a bureaucrat regarding a decision made by Parliament. In contrast, an authoritative text is deemed to have been written by the person or body that enacted it. Of course, in reality such texts are drafted by bureaucrats or lawyers. But courts regard the words in a modern will as those of the testator, just as they consider the legislature as being the author of a statute. This attribution of authorship is obviously a large part of the reason that legal texts of this kind are viewed as so authoritative.

The Creation of Legal Texts

The Enactment or Execution of a Text

Legal texts, especially authoritative texts, tend to be created or executed in relatively formal ways. Wills are a good example. Once a will has been drafted, the testator typically gathers with two witnesses in a room. The testator declares that this is his will and signs it in the presence of the witnesses. The witnesses, who must see the testator sign or acknowledge the will, then each sign the will. Although the rules are less strict than they once were, defects in the execution process can still be fatal. A typed will that is signed by the testator and notarized is usually not valid. An important purpose of the formalities is to help us identify this document as an authoritative legal text.

The situation is different with contracts, which as we have observed can be oral and relatively informal. Certain types of contractual transactions must be evidenced by a writing, but most need not be. Even when the provisions of an agreement are reduced to writing, the signing of the writing can occur without any particular formalities (although a handshake or drinking a glass of champagne may be customary in some situations).

The lower level of formality in contract law does have a drawback: It may be unclear whether the parties have made a contract or not. Many contract cases revolve around whether the parties made an agreement that is legally enforceable or whether they were merely involved in preliminary discussions or negotiations. In contrast, when someone has gone through the formalities of executing a will, there is little doubt that she intended to make a legally valid will.

In the area of public law, statutes – which, like wills, are highly textualized – usually are enacted with great solemnity and according to established procedures. Since the early 1400s, it has been established in England that a statute must be debated and accepted by both houses, Lords and Commons, and then receive the assent of the monarch. Most American jurisdictions have similar requirements embodied in their constitutions. Promulgation or publication may also be required. The purpose of these formal procedures is to guarantee that a piece of paper or parchment does, in fact, constitute a particular type of authoritative legal text, a statute.

What Does the Text Include?

Once a legal document has been formally executed or enacted, it would seem to be a trivial matter to determine what is part of the text and what is not. It is certainly true that the formalities of executing a text such as a statute usually make the content of the statute sufficiently clear. Only text that has gone through the formal enactment process is part of the relevant statute. Explanatory materials or reports by legislative committees are not part of the text, although they may sometimes be helpful in resolving ambiguities.

A consequence is that additions and changes to the text must be made in an equally formal manner. A legislature cannot delete a sentence in a statute by ordering the clerk to cross it out on the official copy, even if a majority votes in favor of the idea and it obtains presidential approval or royal assent. Rather, the legislature must enact a new statute, following all the requisite formalities, that declares that the previous statute is to be amended by striking out the sentence.

Formalities required for textual modification of private documents are usually somewhat less strict because they were originally enacted or executed

with fewer formalities. There are a number of exceptions, of course, such as written contracts and wills. Like a statute, a will can be modified only by a formal amendment, called a codicil, which must be executed following the same procedures as the will itself.

Whereas the legislature may be expected to know how to properly amend a statute, ordinary citizens are far less aware of the textual conventions of the law when it comes to private legal transactions. Suppose that Uncle Bob has a provision in his will giving $5000 to Alice, his favorite niece. A year before his death he crosses out '$5000' and writes above it '$10 000.' What does Alice get? In many common-law jurisdictions, the interlineation is ineffective, and Alice gets the original sum of $5000. In others, the act of crossing out the words or numbers will be effective, but the insertion of additional text (the number '$10 000') was not done with proper formalities and therefore is invalid; Alice gets nothing. Virtually no jurisdiction would give Alice the full $10 000 that Uncle Bob intended to give her. The reason, of course, is that a person cannot change an authoritative text by such informal means. To modify a will, you have to execute a new will or a codicil with all the required formalities.

In recent years, courts have tried to mitigate the harshness of these textual conventions in various ways, especially when ordinary people with no legal training were involved. Even with such reforms, however, it is clear that the textual conventions of the law are quite different from those relating to ordinary speech and writing.

Expanding the Text by Incorporation

Even though the rules relating to the formation and revision of legal texts can be quite rigid, the concept of text is in some ways surprisingly abstract. For instance, we have also seen that one way to add to the text is by means of an amendment or codicil. An interesting aspect of the process of amendment is that legally the result is that there is now one new text – even though that text is physically contained in two documents. Thus, a will executed in 1980 and a codicil executed in 1990 are considered a single text. In many jurisdictions, a principle called republication by codicil has the further effect of updating the will, so that all of the text is now considered to have been executed in 1990, even if part of it was actually written in 1980.

Another way to add to the text is by means of what is called incorporation by reference. This is a procedure that allows a document outside of the will itself to be made part of the testamentary text. Thus, if Aunt Hattie adds a clause to her will stating that she incorporates by reference a list of personal property located in her safe deposit box, the list becomes part of the text of her will.

Thus, despite the rigid rules relating to how it is created and given effect, the legal text is actually a rather abstract concept. A testator can execute a will in 1980, make a list of gifts of personal property in 1985, and execute a codicil in 1990 that makes some changes to the 1980 will and incorporates the list by reference. All of these documents, molded or integrated together, constitute a will dated 1990.

Structure and Language of Legal Texts

Structure

The most salient feature of the structure of legal texts is that they are highly formulaic or stereotypical. Some texts can be quite elaborate in terms of structure, of course, but routine legal documents tend to follow a predetermined structure that changes little over time. A statute, for instance, normally has some or all of the following elements:

- Long title

  An Act to establish a Board having the function of promoting and facilitating the use of the Welsh language . . .

- Enactment clause (sometimes omitted in printed or codified versions)

  Be it enacted by the Queen's most Excellent Majesty, by and with the advice and consent of the Lords Spiritual and Temporal, and Commons, in present Parliament assembled, and by the authority of the same, as follows . . .

- Substantive provisions

  (1) The Lord Chancellor may make rules as to the provision and employment of interpreters of the Welsh and English languages for the purposes of proceedings before courts in Wales.

  (2) The interpreters shall be paid, out of the same fund as the expenses of the court are payable, such remuneration in respect of their services as the Lord Chancellor may determine.

  (3) The Lord Chancellor's powers under this section shall be exercised with the consent of the Treasury.

- Exceptions or provisos

  This section shall not apply in relation to a provision which confers, or gives power to confer, a name on any body, office or place . . .

- Definitions

> In this Part of this Act "public body" means a county council, . . . district council or community council . . .

- Short title or citation

> This Act may be cited as the Welsh Language Act 1993.

Possible additional clauses, which typically come at the end of a statute, may relate to matters such as the administration of the act, enforcement, transitional provisions, repeals of earlier acts, and schedules.

Private legal documents such as contracts also tend to have a relatively fixed structure. Because there are a large number of different types of contracts, however, there are many different formulas or templates in use. Wills, as might be expected, follow a more routine structure, because they virtually all have the same purpose: the transfer of property at death, as well as some ancillary functions such as appointing an executor or a guardian for minor children. A typical will almost always bears the caption 'Last Will and Testament.' An introductory paragraph typically identifies the testator and revokes any previous wills. It is followed by bequests (gifts of property or amounts of money). Most wills begin with a list of specific bequests, followed by a gift of what is called the residue (all the rest of the estate). Additional clauses may deal with the appointment of an executor or a guardian of minor children. The will normally ends with the signature of the testator and two witnesses.

Wills and statutes are but two examples of the archaic and stereotypical structure of many legal texts. One of the main reasons for this repetitious structure and language is that the lawyers who draft these wills tend to base their work on forms that they have inherited from older lawyers or that they find in books of forms. Not just the structure but the language itself tends to be ritualistic and archaic, as shown in the next section.

Language

Legal language, broadly construed as the language of the legal profession, has been the object of numerous studies, many of which advocate reforming it to make it more understandable to the ordinary citizens whose lives and fortunes may be affected by it. Many pejorative adjectives have been used to describe the language of the law, including 'wordy,' 'unclear,' 'pompous,' and 'dull.' More specifically, the literature on legal language suggests that it differs in a number of ways from ordinary speech, including the use or overuse of technical terminology; archaic, formal, and unusual or difficult vocabulary; impersonal constructions; nominalizations, passive constructions (*the girl was injured* . . .); multiple negation; long and complex sentences; and wordiness and redundancy. Collectively, these features are often called legalese.

It is interesting to observe that when the literature refers to features of legal language, it is almost always in reference to written legal language. Recent studies of courtroom interaction or conversations between lawyers and clients reveal that the members of the legal profession normally do not speak legalese. A lawyer arguing to a jury, discussing a matter with a client, or even arguing a point of law to a judge wishes to be understood and generally employs formal but fairly ordinary standard English. Lawyers and judges inevitably use a lot of technical vocabulary, but their speech usually does not contain an inordinate amount of legalese.

In reality, many or most of the commonly mentioned characteristics of legalese are not so much features of legal language *per se* as they are features of legal texts. In fact, it is generally true that the most authoritative texts (statutes and wills) tend to have the highest percentage of legalese in them. To some extent, legalese – and thus, the language of legal texts – can be considered to be an extreme version of formal written language. Research by linguists into the differences between speech and writing has pointed out that, as a general matter, written text is more planned and speech more spontaneous. Related to this difference is that a hearer must process an utterance as it is being spoken, whereas a reader can go over a written sentence again and again until she has figured it out. Such differences are easy to exaggerate, but it certainly is true that a legal text is planned in advance, which makes it possible for sentences to be much longer, denser, and more complex than speech. Written language has also been correlated to greater occurrences of nominalized verbs and passive constructions, both of which are also considered to be features of legalese.

Although both the structure and the language of legal texts tend to remain formulaic and archaic, there is a growing movement that favors making these texts more accessible to both the public and the legal profession. Particularly in some Australian states, parliamentary drafters have experimented with the structure of statutes, adding flowcharts and examples to help readers understand the content. In various parts of the world, there has been a movement to make the language of legal texts – especially those that affect the welfare of ordinary consumers – more comprehensible. There is much still to be accomplished, however.

Interpreting Legal Text

One of the most often-observed features of legal interpretation is the tendency of many judges and lawyers to interpret legal texts in relatively literal or acontextual ways. For instance, English judges during the 18th and 19th centuries developed the Plain Meaning Rule, under which an interpreting judge was supposed to look at the text of a legal document and determine whether its meaning was 'plain,' based purely on the text contained within the four corners of the document. If so, the meaning of the document should be derived solely from the text itself, without reference to any 'extrinsic evidence' (outside sources). Thus, the meaning of a statute could not take into account statements made by legislators during debates or reports by legislative committees. Only if the meaning of the text was ambiguous could judges appeal to outside sources to help determine its meaning.

The Plain Meaning Rule held sway in England throughout the 20th century, although it has recently lost some of its force. In the United States, it came under attack by both judges and legal academics over the course of the previous century. Many courts began to look at the outside evidence to determine whether the language of a text really was as plain as it seemed. If the external evidence revealed an ambiguity, they would then consider the evidence in determining the meaning of the text. In the area of statutory interpretation, it became relatively acceptable to look at a statute's legislative history to help determine the meaning of an ambiguous word or phrase. Sometimes a judge's view of the intent of the legislature, or of the purpose the statute, might even justify going beyond what seemed to be its 'plain' meaning.

At the end of the 20th century a counterrevolution arrived in the form of what is now called 'textualism.' Its strongest advocate is Justice Antonin Scalia of the U.S. Supreme Court. Essentially, textualists urge judges to focus more on the text of a statute and less on the intentions of the legislature. In that sense, textualism could be viewed as attempting to resuscitate the Plain Meaning Rule.

Those who oppose textualism tend to place greater emphasis on the intentions of the legislature (intentionalists) or on the goal or purpose of a statute (purposivists). They often make the linguistic argument that textualism rests on a misguided or simplistic view of language. No text can have a meaning independent of the intentions of its maker. In fact, the search for meaning invariably focuses on what a speaker meant by an utterance.

It is certainly true that in ordinary spoken language, we normally aim to determine the meaning that a speaker intended to convey and use any pragmatic information that might be available to determine the speaker's communicative intentions. Whether we inevitably use the same strategy to interpret written language and, in particular, authoritative legal texts is less clear.

One important difference between spoken language and written text is that writing tends to be more permanent than speech. This is certainly true of legal texts, which on occasion remain in force for hundreds of years. Furthermore, writing tends to be more planned, whereas speech is usually relatively spontaneous. This is also true of legal drafters, who plan texts carefully in an effort to cover every base and foresee every future contingency. They will never completely succeed, of course, but the fact that they try is in itself significant.

A related difference is that a speaker is usually engaged in face-to-face contact with the hearer, whereas a writer normally is not. Consequently, it is possible for a speaker to monitor whether the hearer seems to understand – and, if not, to explain again in different words. Written legal texts usually do not allow for such subsequent explanation and elaboration. Thus, the drafter must take time to be as clear and precise as possible because he will not have a second chance to explain something. This is why legal drafters tend to be so obsessed with precision. Whether they achieve it is another issue, but judges who interpret a text assume that the drafter chose his words carefully. If a drafter of a statute uses two words that are similar in meaning, such as *residence* and *domicile*, the assumption will be that this is not free variation but that two distinct meanings were intended.

Moreover, spoken interaction often involves people who know something about one another or who can infer certain information about their interlocutor based on the situation in which they find themselves. Face-to-face interaction also allows for the communication of additional information by nonverbal means. The hearer may thus use, in addition to language, shared background information or nonverbal cues to determine the communicative intentions of the speaker. Writers must normally signal all their intentions verbally. This difference promotes a somewhat more literal interpretation of written texts. An utterance spoken to a friend might well be understood, as intended, to be a joke or ironic comment based on what the friend knows about the speaker and the speaker's tone of voice. Someone who read the same words on paper might take them more literally.

Because writing is relatively permanent, legal drafters can create texts that might have to be interpreted

many years or even centuries in the future by an unknown audience. For the most part, all that the drafters have at their disposal to accomplish this aim is the written word. In terms of background information, the main thing drafters can assume about their audience is a basic knowledge of the world and the legal system. Such texts will therefore have to be written in a highly autonomous style, in which the drafter aims to place into the text as much information as is needed to interpret it. The ideal is that a legally educated reader will be able to interpret the text without recourse to other information, even if situated thousands of miles away or many years later. Lawyers are deemed to place all relevant intentions into the text, and judges who interpret the text presume that they were successful. Critically, this ideal provides the underpinnings of the Plain Meaning Rule.

Of course, this idealization of how legal texts are created and interpreted flies in the face of what we know about human language. Language is not always clear, and there is no such thing as a fully autonomous text. Most members of the legal profession are aware of these limitations, which explains why even judges who adhere to the Plain Meaning Rule do not apply it when the text is ambiguous. Nonetheless, judges have often had an inflated notion of how the textual conventions of the profession enable judges to interpret a carefully drafted text without recourse to outside materials.

Even if the profession's textual conventions facilitate or encourage a relatively literal mode of interpretation, it does not follow that a textualist approach to interpreting legal texts is inevitable. A textual approach effectively privileges the text over authorial intent. An intentionalist approach, on the other hand, pays greater attention to the intentions of the drafter or legislature. Which approach is 'best' is ultimately a political decision that depends largely on the current view of the institutional role of lawmaker and judge. In a democratic system in which lawmakers are elected by popular vote and are deemed to express the will of the people, it makes sense for judges (who are often not elected) to carry out as best they can the evident intentions of the legislature.

Yet at times it may be more important to preserve the integrity or cohesiveness of a legal text. Criminal law is perhaps the best example. In this area, the rule of law requires that a statute clearly inform citizens of which types of conduct will subject them to punishment. It should be possible for a citizen to know, based on reading a statute, that certain conduct is prohibited. Such a citizen should not be forced to hire a lawyer to investigate whether there might be some legislative history buried in an archive that could potentially enlarge the statute's scope. The nature of the audience should also be considered. If a legal document such as a contract or will is produced by a person who is untrained in the law and who is therefore unfamiliar with its textual conventions, there is no logical reason to interpret that document in a textual way.

Ultimately, interpretation involves a symbiotic relationship between the encoder and decoder of language. To the extent that language is encoded using particular textual conventions, it should be decoded taking those conventions into account. If, on the other hand, it is encoded as ordinary language, it should be decoded as ordinary speech or writing.

Conclusion

It has long been known that the language of the legal profession is different from ordinary speech. To some extent, these differences can be accounted for by the fact that legal documents tend to be drafted in formal literary English. Many of the features of legal language are also found in the more formal written varieties of ordinary English.

Yet there are also some important ways in which legal texts are quite distinct from other writings. This results in large part from the textual conventions of the profession, which enable the creation of a text that, once properly enacted or executed, comes to be viewed as the authoritative expression of the intentions of its author. The formal creation of such texts tends to promote a relatively textual, or literal, interpretation.

The formalities surrounding the creation and interpretation of authoritative legal texts can provide a degree of certainty to their makers and users. But those same formalities, because they differ from how ordinary people tend to write and understand documents, can also frustrate the intentions of legislators, testators, and others deemed to be the authors of the text in question. In all areas of the law, there has been a tendency in recent years to draft and enact legal texts in less formal ways and to interpret them in ways that are more consistent with the intentions of their makers. With the exception of some areas of the law where the authoritative text plays a particularly important function, this trend seems likely to continue.

Less clear is the impact that new technologies will have on legal conventions relating to text. It is now possible to embed hyperlinks containing definitions or cross-references to a document. And text no longer necessarily consists of ink marks on paper. How this will change the legal conception of text is unclear, but it seems inevitable that it will.

See also: Law and Language: Overview; Performative Clauses; Speech Act Verbs.

Bibliography

Baker J H (1990). *An introduction to English legal history* (3rd edn.). London: Butterworths.

Baron N (2000). *Alphabet to email: how written English evolved and where it's heading.* London and New York: Routledge.

Bowers F (1989). *Linguistic aspects of legislative expression.* Vancouver, Canada: University of British Columbia Press.

Chafe W L (1982). 'Integration and involvement in speaking, writing, and oral literature.' In Tannen D (ed.) *Spoken and written language.* Norwood, NJ: Ablex. 35–53.

Danet B & Bogoch B (1994). 'Orality, literacy and performativity in Anglo-Saxon wills.' In Gibbons J (ed.) *Language and the law.* London and New York: Longman. 100–135.

Farnsworth E A (1990). *Contracts* (2nd edn.). Boston: Little, Brown & Co.

Hiltunen R (1990). *Chapters on legal English: aspects past and present of the language of the law.* Helsinki: Suomalainen Tiedeakatemia.

Kay P (1977). 'Language evolution and speech style.' In Blount B G & Sanches M (eds.) *Sociocultural dimensions of language change.* New York: Academic Press. 21–33.

McGovern W M *et al.* (1988). *Wills, trusts and estates* (3rd edn.). St. Paul, MN: West Publishing.

Mellinkoff D (1963). *The language of the law.* Boston: Little, Brown & Co.

Murray J E (1990). *Murray on contracts.* Charlottesville, VA: The Michie Co.

Plucknett T F T (1956). *A concise history of the common law.* Boston: Little, Brown & Company.

Solan L M (1993). *The language of judges.* Chicago, IL: University of Chicago Press.

Tiersma P M (1999). *Legal language.* Chicago, IL: University of Chicago Press.

Tiersma P M (2001). 'Textualizing the law.' *Forensic Linguistics 8(2),* 73–92.

Whitelock D (1930). *Anglo-Saxon wills.* Cambridge, UK: Cambridge University Press.

Language of Tertiary Education

U Ammon, Universitaet Duisburg-Essen, Duisburg, Germany

Languages as Medium of Tertiary-Level Teaching: Overview

Of the some 6000 languages presently spoken worldwide ('living' languages), only a small fraction serve as the medium of teaching on the tertiary level, i.e., for tertiary-level teaching. More are used for secondary-level teaching, and a still larger portion for primary-level teaching. The number of languages that are studied as teaching subjects (as opposed to medium) is even greater, with probably larger numbers on the tertiary than on the lower educational levels. The teaching subjects comprise many of the living as well as some of the 'dead' languages, if one includes teaching for small groups of linguistic experts around the world. For a language to function as a tertiary-level medium requires its sufficient modernization (*Ausbau* – cf. Kloss, 1978: 37–60), especially availability of adequate terminology and texts, which in turn depends on the language's history and its community's technical advancement, economic strength and linguistic resolution.

A reasonably reliable source (Salzman, 2002) mentions, in its extensive (though incomplete) country-by-country list of tertiary institutions, 82 languages altogether for teaching on the tertiary level, of which 39 are used in more than one country. **Table 1** lists those 13 languages that are used in three or more countries with the number of countries given in each case. None of them is the sole tertiary medium in each single one of these countries but sometimes used only for special subjects or study programs. However, no less than 41 languages, i.e., half of the entire list, appear to be the sole tertiary medium in at least one country; but figures have to be taken with caution because of the incompleteness of data. The rest, again 41 languages, function only as the co-medium, presumably only for a limited range of scholarly or scientific fields. Some of the languages excluded from may, however, be used more regularly and in more scientific or scholarly fields than some of those listed in the table.

Official Language and Medium of Teaching

All of the 82 languages functioning as the medium of teaching on the tertiary level are the official language of some country, mostly even the national official

See also: Law and Language: Overview; Performative Clauses; Speech Act Verbs.

Bibliography

Baker J H (1990). *An introduction to English legal history* (3rd edn.). London: Butterworths.

Baron N (2000). *Alphabet to email: how written English evolved and where it's heading*. London and New York: Routledge.

Bowers F (1989). *Linguistic aspects of legislative expression*. Vancouver, Canada: University of British Columbia Press.

Chafe W L (1982). 'Integration and involvement in speaking, writing, and oral literature.' In Tannen D (ed.) *Spoken and written language*. Norwood, NJ: Ablex. 35–53.

Danet B & Bogoch B (1994). 'Orality, literacy and performativity in Anglo-Saxon wills.' In Gibbons J (ed.) *Language and the law*. London and New York: Longman. 100–135.

Farnsworth E A (1990). *Contracts* (2nd edn.). Boston: Little, Brown & Co.

Hiltunen R (1990). *Chapters on legal English: aspects past and present of the language of the law*. Helsinki: Suomalainen Tiedeakatemia.

Kay P (1977). 'Language evolution and speech style.' In Blount B G & Sanches M (eds.) *Sociocultural dimensions of language change*. New York: Academic Press. 21–33.

McGovern W M *et al.* (1988). *Wills, trusts and estates* (3rd edn.). St. Paul, MN: West Publishing.

Mellinkoff D (1963). *The language of the law*. Boston: Little, Brown & Co.

Murray J E (1990). *Murray on contracts*. Charlottesville, VA: The Michie Co.

Plucknett T F T (1956). *A concise history of the common law*. Boston: Little, Brown & Company.

Solan L M (1993). *The language of judges*. Chicago, IL: University of Chicago Press.

Tiersma P M (1999). *Legal language*. Chicago, IL: University of Chicago Press.

Tiersma P M (2001). 'Textualizing the law.' *Forensic Linguistics 8(2)*, 73–92.

Whitelock D (1930). *Anglo-Saxon wills*. Cambridge, UK: Cambridge University Press.

Language of Tertiary Education

U Ammon, Universitaet Duisburg-Essen, Duisburg, Germany

Languages as Medium of Tertiary-Level Teaching: Overview

Of the some 6000 languages presently spoken worldwide ('living' languages), only a small fraction serve as the medium of teaching on the tertiary level, i.e., for tertiary-level teaching. More are used for secondary-level teaching, and a still larger portion for primary-level teaching. The number of languages that are studied as teaching subjects (as opposed to medium) is even greater, with probably larger numbers on the tertiary than on the lower educational levels. The teaching subjects comprise many of the living as well as some of the 'dead' languages, if one includes teaching for small groups of linguistic experts around the world. For a language to function as a tertiary-level medium requires its sufficient modernization (*Ausbau* – cf. Kloss, 1978: 37–60), especially availability of adequate terminology and texts, which in turn depends on the language's history and its community's technical advancement, economic strength and linguistic resolution.

A reasonably reliable source (Salzman, 2002) mentions, in its extensive (though incomplete) country-by-country list of tertiary institutions, 82 languages altogether for teaching on the tertiary level, of which 39 are used in more than one country. Table 1 lists those 13 languages that are used in three or more countries with the number of countries given in each case. None of them is the sole tertiary medium in each single one of these countries but sometimes used only for special subjects or study programs. However, no less than 41 languages, i.e., half of the entire list, appear to be the sole tertiary medium in at least one country; but figures have to be taken with caution because of the incompleteness of data. The rest, again 41 languages, function only as the co-medium, presumably only for a limited range of scholarly or scientific fields. Some of the languages excluded from may, however, be used more regularly and in more scientific or scholarly fields than some of those listed in the table.

Official Language and Medium of Teaching

All of the 82 languages functioning as the medium of teaching on the tertiary level are the official language of some country, mostly even the national official

language. A few are only regional official languages, notably in India (Assamese, Dogri, Gujarati, Kannada, Malayalam, Marathi, Oriya, Punjabi, Sindhi, Telugu), Russia (Chuvash, Komi, Tatar, Yakut), and Spain (Basque, Catalan [also Valencian], Galician). Not all the national official languages of countries function as the tertiary-level teaching medium, not to speak of the regional official languages. Those excluded are typically the languages of small communities on the one hand, whose members are, at the same time, bilingual in the language of a larger community, e.g., not Luxemburgish (Luxemburg) or Romansch (Switzerland). Thus, the speakers of Luxemburgish or Romansch are at the same time bilingual in French or German, through which they have access to tertiary education. A somewhat different case, on the other hand, includes the indigenous national languages of former colonial countries, which play no or only a marginal role alongside the former colonial languages which remain the main tertiary teaching medium, especially in anglophone and francophone Africa (cf. Ridge, 1999), but partially also in the Arabic-speaking countries (cf. Suleiman, 1999) and some Asian countries (for Malaysia, e.g., Kaur Gill, 2002: 103–118). It seems, however, safe to assume, that all nationally or regionally official languages are subjects of teaching (at least subjects of choice) on the tertiary level of their countries.

Differences According to Disciplines and Levels of Teaching

It is a well-known fact that the number of languages for teaching natural sciences is more limited than for teaching the humanities (cf. Kloss, 1978: 46–55). Especially for graduate science courses, comprehensive teaching materials are only available in a small number of languages, while oral teaching extends to more languages. Rigorously speaking, English may even be the only 'fully equipped' language today. From a less rigorous viewpoint, the languages for which reasonably comprehensive graduate science teaching materials are available are roughly those listed in Table 1, with some of them less equipped than others (Armenian, Arabic, Turkish) and a few even better equipped than some of those included in the table (e.g., Hebrew, Hungarian, Polish, Swedish). Advanced science teaching materials can, as a rule, only be produced by language communities, or countries, who themselves play a notable role in science research and who have a tradition to do so as well as an adequately modernized language. Otherwise, they have to make do with a 'foreign' language.

Table 1 Languages used for tertiary-level teaching in three or more countries

Language	*Number of countries*
English	103
French	42
Arabic	23
Spanish	18
German	15
Russian	10
Dutch	6
Portuguese	5
Chinese	5
Japanese	3
Italian	3
Armenian	3
Turkish	3

In the humanities and, to some extent, in the social sciences, a greater number of languages are employed for tertiary-level teaching than in the natural sciences. Here it is more likely for language communities, or countries, to use their own nationally or regionally official languages for various reasons: topics seem to be more immediately relevant for their own national identity and their own language community, and a respectable level of research seems to be easier, or less costly, to achieve than in the 'hard' sciences, which is why advanced teaching can interact quite directly with their ongoing own research.

Recent Changes in the Role of English

Until a couple of decennia ago, countries such as France, Germany, Italy, and Japan offered all tertiary-level teaching, even in the natural sciences, in their own languages. This has, however, changed in recent years, with English being introduced as an additional language in all of these countries at least at some institutions and for special programs (cf. Ammon and McConnell, 2002, for European countries). It seems safe to say that nowadays all the non-English-speaking countries make at least some use of English for tertiary-level teaching, especially in the natural sciences and modern technologies but also, to a lesser extent, in the social sciences including economics. There are several reasons for this spread of English. One is intensification of global science communication that favors the use of one single language, naturally the one most widespread and used by the scientifically leading country (the United States). Most countries are eager to equip their own academics and students with skills in this language, which appear indispensable for successful global

communication (cf. Crystal, 1997). Another reason for the introduction of English-medium teaching on the tertiary level is competition on the international academic market, especially the craving for foreign students and scientists or scholars (Van Parijs, 2000). The English-speaking countries reap huge benefits from the preference of internationally mobile academics, not only in terms of money (e.g., tuition fees; cf. Kaplan, 2001) but also in terms of international connections including spread of own values and ideology. A major reason why they win the lion's share in this market is the English language, in which most academics or students have acquired skills in school and which they would like to further improve rather than acquiring skills in another foreign language. Countries such as France or Germany have reasons to believe that insistence on their own language would keep foreign academics and students out, which is why they started to introduce English for tertiary-level teaching. Their scientists and, to a lesser extent, their scholars have shifted to English as their preferred language of publication years ago (cf. Ammon, 1998).

It should be stressed that none of these countries plans to completely shift to English for tertiary-level teaching. On the contrary, official policies aim at containing English at the level of merely an additional language and at maintaining the local language as the primary teaching medium. Otherwise, fears are, foreign language studies and prestige of the local tongue would suffer. The possibility of studying in France or Germany still is considered an important motive for studying French or German as a foreign language. On the face of it, official policies try to prevent English from completely replacing the local language and becoming the sole tertiary teaching medium. These policies aim at limiting the function of English to improving the local academics' and students' skills in the language and to opening up tertiary institutions for English-speaking foreigners. Once the latter have entered, they are expected to learn the local language, too, and, in the course of time, to do part of their studies in it. As long as skills in the local language remain a requirement for successful studies, it appears worthwhile studying the language abroad, before entering the country, to avoid the burden of extra language learning alongside regular studies. Thus, negative effects of introducing English as a tertiary teaching medium on the departments of French or German abroad could perhaps be kept at bay. As much as such attempts appear in line with the country's own interests, officials are shy to rigorously insist on (for fear of being criticized of national selfishness), which is one of the reasons, among perhaps weightier others, why programs with English as the sole teaching medium tend to spread.

Problems and Attempts at Alleviation

Countries such as France, Germany, and Japan fear that the use of English as their tertiary teaching medium might accelerate decline of studies of their own language and, thus, to loose an important channel of their international relationships. They also tend to see it as a threat to their national identity, for which their own language is an important symbol. As a countermeasure they try, besides stabilizing their own language at home (cf. section 4), to establish own-language study programs abroad. Successful endeavors of this kind partially explain why languages such as French, German or Japanese function as the tertiary-teaching medium beyond the countries where they have official status. English often still remains, in fact, the preferred medium of teaching at tertiary institutes financed by non-English-speaking countries abroad, which have to be content with teaching their own language alongside English-medium studies, as, for example, at the German University in Cairo.

Whether introducing English as a tertiary-level teaching medium has, in the long run, more advantages or disadvantages for countries such as France, Germany, and Japan, is a question difficult to answer. The resulting better skills in English could very well help to break, in the long run, the often bewailed Anglo-Saxon monopoly of access to the most advanced scientific knowledge and publication channels (cf., e.g., Durand, 2001). A thorough success seems, however, only possible if the non-English-speaking countries could establish their forms of English, or 'Globalish' (cf. Ammon, 2003), as linguistically equivalent to native-speaker English. The home–school language gap (cf., e.g., MacLure, 1999) is usually less of a serious problem on tertiary than on lower educational levels, as it can be softened through preparatory language teaching at the lower levels.

Foreign Language Requirements

A final remark seems in order on foreign language requirements for tertiary-level studies. They have, or are believed to have, various functions, a particularly important one of which is providing access to the literature or culture relevant for the subject of study. Thus, students of European or of Asian history are often required to study Latin or Greek or, respectively, Chinese or Sanskrit. Also, science students worldwide are required to study English nowadays. Conversely and conspicuously, foreign language requirements

were abolished at American universities in the 1960s (controversially discussed in several contributions to *Science*, e.g., Ross and Shilling, 1966; Hartmann, 1967), although abolition was not complete: foreign language studies keep being required for certain subjects at most universities. But Ph.D.s in science including medicine can now be acquired without the knowledge of any foreign language at many American universities. This of course forces scientists of another linguistic background to use English for contacts with their American colleagues. It is, finally, the unequal power of language communities that determines language choice in science, in research as well as in teaching.

See also: Education in a Former Colonial Language; Foreign Language Teaching Policy; Language Policies: Policies on Language in Europe; Language Spread; Languages of Wider Communication; Lingua Francas as Second Languages; Pragmatics: Linguistic Imperialism; Second and Foreign Language Learning and Teaching.

Bibliography

Ammon U (1998). *Ist Deutsch noch internationale Wissenschaftssprache? Englisch auch für die Lehre an deutschsprachigen Hochschulen.* Berlin & New York: Walter de Gruyter.

Ammon U (2003). 'Global English and the non-native speaker: overcoming disadvantage.' In Tonkin H & Reagan T (eds.) *Language in the 21st century.* Amsterdam: John Benjamins. 23–34.

Ammon U & McConnell G (2002). *English as an academic language in Europe: a survey of its use in teaching.* Frankfurt: Peter Lang.

Crystal D (1997). *English as a global language.* Cambridge: Cambridge University Press.

Durand C (2001). *La Mise en place de monopoles du savoir.* Paris: L'Harmattan.

Hartmann J D (1967). 'Language requirements modified.' *Science 155*, 626.

Kaplan R (2001). 'English – the accidental language of science?' In Ammon U (ed.) *The dominance of English as a language of science.* Berlin & New York: Mouton de Gruyter. 3–26.

Kaur Gill S (2002). *International communication: English language challenges for Malaysia.* Serdang: Universiti Putra Malaysia Press.

Kloss H (1978). *Die Entwicklung neuer germanischer Kultursprachen seit 1800.* Düsseldorf: Schwann.

MacLure M (1999). 'Home and School Language.' In Spolsky B (ed.) *Concise encyclopedia of educational linguistics.* Oxford, UK: Elsevier. 202–204.

Ridge S G M (1999). 'Language education policy – Africa.' In Spolsky B (ed.) *Concise encyclopedia of educational linguistics.* Oxford: Elsevier. 101–106.

Ross S & Shilling C W (1966). 'Language Requirements for the Ph.D.' *Science 153,* 1595.

Salzman M (ed.) (2003). *The world of learning.* London/New York: Europa Publications.

Suleiman Y (1999). 'Language education policy – Arabic speaking countries.' In Spolsky B (ed.) *Concise encyclopedia of educational linguistics.* Oxford, UK: Elsevier. 106–116.

Van Parijs P (2000). 'The ground floor of the world. On the socio-economic consequences of linguistic globalisation.' *International Political Science Review 21,* 217–233.

Language of Thought

R Carston, University College London, London, UK

The idea that there is a language of thought amounts to the following: having a thought with a particular content is a matter of being related in a certain way to a sentence in an innately given mental language. The sentences or formulas in this language (Mentalese) are like the sentences of public natural languages (Japanese, Urdu, English, etc.) in that they have syntactic and semantic properties. That is, they are composed of constituents in particular structural configurations and their semantics is a function of the semantics of their basic elements and their syntactic structure. The grammar of Mentalese may, however, differ quite radically from the grammar of any natural language. The basic elements (concepts, perhaps) denote entities and properties in the world. The full formulas are truth-conditional, and so have truth values as determined by the way the world is, and they bear logical relations to each other, such as entailment.

The Representational Theory of Mind

The relevant notion of a thought here is that of an 'intentional' state of mind, where intentional mental states are those that have the property of being representational, that is, of being about the world. Beliefs, desires, intentions, hopes, and fears are different types of intentional mental states. They are sometimes

were abolished at American universities in the 1960s (controversially discussed in several contributions to *Science*, e.g., Ross and Shilling, 1966; Hartmann, 1967), although abolition was not complete: foreign language studies keep being required for certain subjects at most universities. But Ph.D.s in science including medicine can now be acquired without the knowledge of any foreign language at many American universities. This of course forces scientists of another linguistic background to use English for contacts with their American colleagues. It is, finally, the unequal power of language communities that determines language choice in science, in research as well as in teaching.

See also: Education in a Former Colonial Language; Foreign Language Teaching Policy; Language Policies: Policies on Language in Europe; Language Spread; Languages of Wider Communication; Lingua Francas as Second Languages; Pragmatics: Linguistic Imperialism; Second and Foreign Language Learning and Teaching.

Bibliography

Ammon U (1998). *Ist Deutsch noch internationale Wissenschaftssprache? Englisch auch für die Lehre an deutschsprachigen Hochschulen.* Berlin & New York: Walter de Gruyter.

Ammon U (2003). 'Global English and the non-native speaker: overcoming disadvantage.' In Tonkin H & Reagan T (eds.) *Language in the 21st century.* Amsterdam: John Benjamins. 23–34.

Ammon U & McConnell G (2002). *English as an academic language in Europe: a survey of its use in teaching.* Frankfurt: Peter Lang.

Crystal D (1997). *English as a global language.* Cambridge: Cambridge University Press.

Durand C (2001). *La Mise en place de monopoles du savoir.* Paris: L'Harmattan.

Hartmann J D (1967). 'Language requirements modified.' *Science 155*, 626.

Kaplan R (2001). 'English – the accidental language of science?' In Ammon U (ed.) *The dominance of English as a language of science.* Berlin & New York: Mouton de Gruyter. 3–26.

Kaur Gill S (2002). *International communication: English language challenges for Malaysia.* Serdang: Universiti Putra Malaysia Press.

Kloss H (1978). *Die Entwicklung neuer germanischer Kultursprachen seit 1800.* Düsseldorf: Schwann.

MacLure M (1999). 'Home and School Language.' In Spolsky B (ed.) *Concise encyclopedia of educational linguistics.* Oxford, UK: Elsevier. 202–204.

Ridge S G M (1999). 'Language education policy – Africa.' In Spolsky B (ed.) *Concise encyclopedia of educational linguistics.* Oxford: Elsevier. 101–106.

Ross S & Shilling C W (1966). 'Language Requirements for the Ph.D.' *Science 153,* 1595.

Salzman M (ed.) (2003). *The world of learning.* London/New York: Europa Publications.

Suleiman Y (1999). 'Language education policy – Arabic speaking countries.' In Spolsky B (ed.) *Concise encyclopedia of educational linguistics.* Oxford, UK: Elsevier. 106–116.

Van Parijs P (2000). 'The ground floor of the world. On the socio-economic consequences of linguistic globalisation.' *International Political Science Review 21,* 217–233.

Language of Thought

R Carston, University College London, London, UK

The idea that there is a language of thought amounts to the following: having a thought with a particular content is a matter of being related in a certain way to a sentence in an innately given mental language. The sentences or formulas in this language (Mentalese) are like the sentences of public natural languages (Japanese, Urdu, English, etc.) in that they have syntactic and semantic properties. That is, they are composed of constituents in particular structural configurations and their semantics is a function of the semantics of their basic elements and their syntactic structure. The grammar of Mentalese may, however, differ quite radically from the grammar of any natural language. The basic elements (concepts, perhaps) denote entities and properties in the world. The full formulas are truth-conditional, and so have truth values as determined by the way the world is, and they bear logical relations to each other, such as entailment.

The Representational Theory of Mind

The relevant notion of a thought here is that of an 'intentional' state of mind, where intentional mental states are those that have the property of being representational, that is, of being about the world. Beliefs, desires, intentions, hopes, and fears are different types of intentional mental states. They are sometimes

called propositional attitudes, since they involve the having of an attitude to a content or proposition, for example, having the belief attitude to the 'Paris is a beautiful city' content. These intentional mental states play a central role in cognitive psychology in the explanation of human purposeful behavior. For example, it is because Madge wants to drink some Pepsi and she believes that there is some Pepsi in her refrigerator that she goes to the refrigerator and reaches inside it. On the language of thought (LoT) view, having a belief or a desire, etc., with a certain content entails being in a relation to an internally represented sentence with that content, so the explanation of Madge's refrigerator-oriented behavior will include a specification of the interaction of the sentences that represent the content of her relevant beliefs and desires.

The LoT hypothesis arises in the context of the current computational model of the mind, whereby mental processes, such as reasoning, are sequences of mental states and the transitions between states are effected computationally. Conceiving of these computations as formal/syntactic operations defined over mental representations gives a mechanical explanation for mental processes. That is, they operate on symbols in virtue of the form of the symbol, not in virtue of any semantic property of the symbol, just like the operations performed by a computer or the transitions from line to line in a logic proof. This approach to the causal explanation of mental processes is known as 'methodological solipsism' (Fodor, 1981; Lycan, 1990a). It follows that two beliefs or desires are cognitively distinct if and only if the representations of their contents are formally distinct. For example, consider the desire to meet the husband of Janet Fodor and the desire to meet the staunchest advocate of the language of thought hypothesis. These are identical in their truth-conditional content (given that the definite description in each case picks out the same individual in the world, namely Jerry Fodor). However, so far as cognitive activity is concerned, these are quite distinct types of desire, as they may each be the result of a distinct sequence of thought, and each may cause further different thoughts. Furthermore, they may issue in quite distinct behaviors: in the first case, one might telephone Janet Fodor to ask her and her husband for dinner; in the second, one might seek out conferences on the philosophy of mind. The crucial point here is that thoughts have their causal roles as a function of their formal properties. Semantic properties are respected only in so far as they are mimicked by formal properties, which, of course, they are to at least some extent, since deductive reasoning, which preserves truth, plays a major role in human thought.

Why Should Thoughts Have Syntactic Structure?

One could be an 'intentional realist,' that is, one could accept (a) that beliefs and desires really exist, (b) that they are physically instantiated in the brain, and (c) that they play a causal role in thinking and acting, without positing a **language** of thought in which the objects of attitudes are couched. What's crucial about language is constituent structure, that is, that a sentence is made up of parts and these same parts can occur in a range of different relations with each other in different sentences. So what distinguishes the LoT view from other intentionally realist views is that it entails that belief/desire states are structured states. Fodor (1975) claimed that the language of thought was implicit in the computational approach to psychological explanation since computation presupposes a medium in which to compute. However, the emergence of an alternative computational approach, connectionism (see Sterelney, 1990 for an introduction and *see* **Associationism and Connectionism**), indicates that more argument for structured thought is required, since according to connectionism the mental causes of intelligent behavior can be modeled by patterns of activation across networks of nodes and connections, involving no level of symbolic representation. One of Fodor's (1987b) arguments for syntactic thought turns on the 'productivity' and 'systematicity' of thought. The set of thoughts is potentially infinite and the ability to think any particular thought is intrinsically connected to the ability to think various other thoughts. So, for example, anyone who can form the thought 'the ruthless spy saw the desperate terrorist' can also form the thoughts 'the desperate terrorist saw the ruthless spy' and 'the desperate spy saw the terrorist,' etc. The parallel with natural language is obvious, and since the explanation for the productivity and systematicity of language is its combinatorial syntax and semantics, it is natural to assume that thought too has combinatorial structure.

The Relation between Thought and Public Language

Fodor believes that the semanticity of natural language, that is, the 'about-the-world' property of linguistic expressions, is dependent on the representationality of thought. So, the answer to the question 'How is it that the English sentence, "Paris is a beautiful city," is about Paris?' is, roughly: 'Because that sentence is a vehicle for expressing a thought about Paris.' On this view, an account of the semanticity of natural language

follows from an account of how thoughts refer to the world. For interesting ideas about this logically prior theory of 'psychosemantics,' see Fodor (1987a) and Millikan (1989).

See also: Associationism and Connectionism; Cognitive Science: Overview; Fodor, Jerry (b. 1935); Philosophy of Mind.

Bibliography

Carruthers P (1996). *Language, thought and consciousness: an essay in philosophical psychology.* Cambridge: Cambridge University Press.

Davies M (1998). 'Language, thought and the language of thought.' In Carruthers P & Boucher J (eds.) *Language and thought: interdisciplinary themes.* Cambridge: Cambridge University Press. 226–247.

Fodor J (1975). *The language of thought.* Sussex: The Harvester Press.

Fodor J (1981). *Representations.* Cambridge, MA: Bradford Books/MIT Press.

Fodor J (1987a). *Psychosemantics: the problem of meaning in the philosophy of mind.* Cambridge, Massachusetts: Bradford Books/MIT Press. Reprinted in Lycan (ed.) (1990b), 282–299.

Fodor J (1987b) 'Why there still has to be a language of thought.' In Fodor (1987a), 135–167. Reprinted in Lycan (ed.) (1990b), 282–299.

Lycan W G (1990a). 'The language of thought hypothesis.' In Lycan (ed.) (1990b). 277–281.

Lycan W G (ed.) (1990b). *Mind and cognition: a reader.* Oxford: Blackwell.

Millikan R (1989). 'Biosemantics.' *The Journal of Philosophy 86(6),* 281–297.

Sterelney K (1990). *The representational theory of mind.* Oxford: Blackwell.

Language Planning and Policy: Models

B Spolsky, Jerusalem, Israel
R D Lambert, University of Pennsylvania, PA, USA

Just as with many new fields, there continues to be disagreement over the name of this one, variously called language policy (Nesiah, 1954; Sibayan, 1974), language treatment (Neustupný, 1970), language cultivation (Prague School, 1973), language engineering (Sibayan, 1974), language planning (Haugen, 1959), and language management (Jernudd, 2001). While the last five are more or less synonymous, referring to attempts by authorities to modify language behavior, the first can refer to the customary practices in choice of language items and variety in a speech community, or to a specific decision or set of decisions to modify those practices. To avoid confusion, we will use the terms as follows. The language policy of a speech community (an undefined term, ranging in size from a family through a nation-state to a multinational grouping) consists of the commonly agreed set of choices of language items – whether sounds or words or grammar – or language varieties – whether codes or dialects or named languages – and the beliefs or ideologies associated with those choices. It can be found in language practices and beliefs or in formal policy decisions such as laws, constitutions, or regulations. Language management, planning, engineering, cultivation, and treatment are actions taken by formal authorities such as governments or other agencies or people who believe that they have authority, such as parents, teachers, or academies, to modify the language choices made by those they claim to have under their control (Spolsky, 2004). Language management itself has three components: the development of explicit language plans and policies, their implementation (by rules or laws or resource allocation), and the evaluation of results and effects (cf. Rubin and Jernudd, 1979: 2–3).

Managing Bad and Good Language

Language policy makers and analysts apply the term policy and its synonyms to a wide variety of administrative levels ranging from international organizations (Van Els, 2001) to world regions (Extra and Gorter, 2001; Kaplan and Baldauf, 2003) to countries (Grenoble, 2003; Lo Bianco and Wickert, 2001), or to single educational institutions (Karyolemou, 2002). The term has been expanded to include what is referred to as 'grass roots language policy', that is, policy originating in or influenced by the affected members of the speech community (Hornberger, 1996). Cooper (1989) shows that it can usefully be applied to a family level.

The main principles of language policy become evident even at this simplest level. In any family, there is language policy, as shown by the normal choices of language in the speech practices of the group, viz., which variety of language is addressed in practice to each member. For instance, baby talk (Ferguson, 1964) may be used with very young children, heritage

follows from an account of how thoughts refer to the world. For interesting ideas about this logically prior theory of 'psychosemantics,' see Fodor (1987a) and Millikan (1989).

See also: Associationism and Connectionism; Cognitive Science: Overview; Fodor, Jerry (b. 1935); Philosophy of Mind.

Bibliography

Carruthers P (1996). *Language, thought and consciousness: an essay in philosophical psychology.* Cambridge: Cambridge University Press.

Davies M (1998). 'Language, thought and the language of thought.' In Carruthers P & Boucher J (eds.) *Language and thought: interdisciplinary themes.* Cambridge: Cambridge University Press. 226–247.

Fodor J (1975). *The language of thought.* Sussex: The Harvester Press.

Fodor J (1981). *Representations.* Cambridge, MA: Bradford Books/MIT Press.

Fodor J (1987a). *Psychosemantics: the problem of meaning in the philosophy of mind.* Cambridge, Massachusetts: Bradford Books/MIT Press. Reprinted in Lycan (ed.) (1990b), 282–299.

Fodor J (1987b) 'Why there still has to be a language of thought.' In Fodor (1987a), 135–167. Reprinted in Lycan (ed.) (1990b), 282–299.

Lycan W G (1990a). 'The language of thought hypothesis.' In Lycan (ed.) (1990b). 277–281.

Lycan W G (ed.) (1990b). *Mind and cognition: a reader.* Oxford: Blackwell.

Millikan R (1989). 'Biosemantics.' *The Journal of Philosophy 86(6),* 281–297.

Sterelney K (1990). *The representational theory of mind.* Oxford: Blackwell.

Language Planning and Policy: Models

B Spolsky, Jerusalem, Israel
R D Lambert, University of Pennsylvania, PA, USA

Just as with many new fields, there continues to be disagreement over the name of this one, variously called language policy (Nesiah, 1954; Sibayan, 1974), language treatment (Neustupný, 1970), language cultivation (Prague School, 1973), language engineering (Sibayan, 1974), language planning (Haugen, 1959), and language management (Jernudd, 2001). While the last five are more or less synonymous, referring to attempts by authorities to modify language behavior, the first can refer to the customary practices in choice of language items and variety in a speech community, or to a specific decision or set of decisions to modify those practices. To avoid confusion, we will use the terms as follows. The language policy of a speech community (an undefined term, ranging in size from a family through a nation-state to a multinational grouping) consists of the commonly agreed set of choices of language items – whether sounds or words or grammar – or language varieties – whether codes or dialects or named languages – and the beliefs or ideologies associated with those choices. It can be found in language practices and beliefs or in formal policy decisions such as laws, constitutions, or regulations. Language management, planning, engineering, cultivation, and treatment are actions taken by formal authorities such as governments or other agencies or people who believe that they have authority, such as parents, teachers, or academies, to modify the language choices made by those they claim to have under their control (Spolsky, 2004). Language management itself has three components: the development of explicit language plans and policies, their implementation (by rules or laws or resource allocation), and the evaluation of results and effects (cf. Rubin and Jernudd, 1979: 2–3).

Managing Bad and Good Language

Language policy makers and analysts apply the term policy and its synonyms to a wide variety of administrative levels ranging from international organizations (Van Els, 2001) to world regions (Extra and Gorter, 2001; Kaplan and Baldauf, 2003) to countries (Grenoble, 2003; Lo Bianco and Wickert, 2001), or to single educational institutions (Karyolemou, 2002). The term has been expanded to include what is referred to as 'grass roots language policy', that is, policy originating in or influenced by the affected members of the speech community (Hornberger, 1996). Cooper (1989) shows that it can usefully be applied to a family level.

The main principles of language policy become evident even at this simplest level. In any family, there is language policy, as shown by the normal choices of language in the speech practices of the group, viz., which variety of language is addressed in practice to each member. For instance, baby talk (Ferguson, 1964) may be used with very young children, heritage

languages (Cummins, 1983) with grandparents, or community languages (Smolicz and Secombe, 1985) with outsiders. In immigrant families, there is commonly a difference in language usage between adults, older children, and younger children. There are socially determined differences concerning who usually has the right to speak and what topics and forms of language are appropriate when the family gathers. In conversations between adult caretakers and children, there are commonly efforts to manage language choice, whether by encouraging the use of one variety or by attempting to discourage bad language. Definitions of what is bad language vary socially. It may consist of presumed mistakes in grammar or pronunciation, or the use of stigmatized forms or expressions such as cursing, obscenity, blasphemy, foreignisms, or inelegancies. Often beliefs or an ideology will be quite different from practice. Immigrant parents may think that their children should use either the heritage or the new language exclusively. Such beliefs may or may not lead to successful efforts at management.

The effort to avoid bad language and teach good language is carried outside the home into other institutions, particularly the school which takes a leading role in efforts to modify the language known and used by its pupils. Because of their central role in language socialization, school teachers are most comfortable, it seems, with a standardized variety of language, with clear statements on what is right and what is wrong. They commonly share the puristic belief, that there is a 'correct' variety of language, and, *contra* King Cnut, the key belief that language management is possible. They believe that they themselves use correct language: French teachers are sure they pronounce the *l* in *il vient* and Palestinian teachers are sure they teach in Modern Standard Arabic.

Correctness, however discovered or defined, is one common criterion for good language (Guitarte and Quintero, 1974). Another is the avoidance of obscenity, sometimes institutionalized in laws and regulations at the local and national government level. The United States has federal laws against obscenity, but standards and definitions are local (Harrison and Gilbert, 2000). Blasphemy, an obvious concern of religious institutions, is unlikely to be a matter of legislation in secular nations, but remains an issue in the many states with religiously dominated constitutions such as Pakistan. Seditious language, as opposed to actual sedition and violent language ('fighting words' in the laws of some southern U.S. states) is also criminalized in some nations.

A more recent criterion for good and bad language is 'political correctness', the avoidance of chauvinist or racist language. The campaign to avoid words or expressions that stigmatize racial or religious groups or that express prejudice based on gender (assigned or constructed) is about half a century old in the West. In the United States it has led to language management efforts especially by publishers and editors who try to ban gender-biased terminology and grammar (Pauwels, 1998).

A common criterion for language management, one that moves us closer to the realm associated with national language policy, is the avoidance of expressions and words considered foreign (Annamalai, 1989). One of the inevitable effects of culture and language contact, and it is difficult to distinguish the two, is a tendency to borrow foreign words along with the new concepts and artifacts that they label. In many situations, ideological opposition to foreign borrowings is nearly as strong as opposition to the use of foreign languages (Kroskrity, 1998). In Latin America from the conquest, there was puristic opposition to the use of 'americanismos', defined as words borrowed from native languages or locally coined, and a similar anti-foreign purism now calls for laws against borrowings from English (Rajagopalan, 2002). The idea of a pure uncontaminated language is widespread (Jernudd and Shapiro, 1989). Preventing linguistic corruption was and remains a key task of the French Academy. Most national language movements hold puristic beliefs, although the particular source of contamination (French in the Dutch-speaking portion of Belgium, Arabic in Turkish, Yiddish in Hebrew, Danish and now English in Icelandic) varies. Presumably, this represents a belief in the identifying and symbolic value of language. By admitting foreign elements, 'I may be weakening my national identity'.

Three important generalizations emerge from the discussion so far. One is the tension between pragmatic communicative goals (for instance, the caretaker aims to give the child the most efficient variety of language) and symbolic and social goals (identifying the speaker with a chosen social group). A second generalization relates to the linguistic and social levels on which policy can apply: linguistically, it can refer to a single sound or word (*ain't*) or to a labeled variety of language (jargon, English); socially, it can apply to a small social group such as a family or to a higher level such as a nation state or an international federation. A third generalization is that policy is manifested in practice, in beliefs or ideology, and in management activities, and while the three aspects are intertwined, they need not be consistent.

National Language Policies

Most analyses of language policy and management are concerned with formal, governmentally backed activities at the national or regional level aimed at

controlling language knowledge and use within a country or region.

During the first phase of the study of language planning, Kloss (1969) proposed a useful distinction between what he called status planning, the determination of the status and functions of a language in a community (such as 'official' or 'national') and what he called corpus planning, the specification of the proper form a particular language should take (such as writing system or spelling or approved lexicon or grammar). Planning was considered an appropriate term in the 1960s, as one part of national development planning (Das Gupta and Ferguson, 1977).

Cooper (1989) added to these two the field of acquisition planning, the determination of which languages should be taught to those who do not speak them and how. While these three domains are conceptually distinct, in practice they overlap. Making a language variety official usually involves standardizing it, writing it down, and modernizing it. It also requires teaching it to citizens who do not know it. Foreign and domestic language policies are blended in situations like the status of French in Canada, the retention of colonial languages in Africa, and the status of trans-border languages in ethnic enclaves such as Swedish in Finland or French and Italian in Switzerland or French and German in Italy.

Corpus Planning and Management

Concern for the form of language may be discerned in the efforts of parents, teachers, and other caretakers to make sure that their charges speak clearly and use forms that are acceptable. It appears more fully developed and institutionalized in efforts to maintain the purity and correctness of sacred texts and in the educational systems that take on some responsibility for correctness. Indeed, language management agencies (Dominguez and López, 1995) are often part of a ministry of education.

More generally, in many countries, especially those where the issue of status is not salient, the largest share of language management is concerned with corpus policy, the prescription of the proper form a language should take, and the cultivation of a language to handle appropriate functions (Prague School, 1973). This can take a variety of forms. In many of the least developed countries and among some indigenous groups in developed countries, the principal corpus activity is the adoption or adaptation of a script and the promotion of literacy among its speakers (Fishman, 1977).

Another frequent goal is language purification (Neustupný, 1989). Commonly this involves an attempt to return to a sometimes fictitious primal language, purging the modern language of loan words and expressions imported from other languages (Jernudd and Shapiro, 1989). Examples are the purging of Persian and the substitution of Sanskrit-based words in Hindi, and the reverse in Urdu (R. King, 2001). Similarly, the deletion of foreign influences in German during the Nazi years and the perpetual struggle of French against Franglish (Weinstein, 1989) are of the same order. Sometimes purification is more extensive. For instance, under Atatürk, a deliberate attempt was made to simplify and modernize Turkish (G. Lewis, 1999). Older linguistic forms borrowed from Ottoman Turkish, Persian, and Arabic were replaced with elements identified with a Turkic past, and the Perso-Arabic script was converted to a Roman one. Similarly, in China the development of Putonghua was accompanied by extensive modernization of vocabulary and morphology (Coulmas, 1991). In a similar vein, the attempt to create a pan-national standard Arabic and to diffuse it throughout the Middle East and North Africa, overlaying the sometimes mutually unintelligible country dialects, has required major innovation in the writing system, grammar, and lexicon of the language (Suleiman, 2003).

Sometimes corpus activity has been directed to the revival or rejuvenation of a language that historically had become fossilized or marginalized, for instance, the attempt to support the use of Quichua in the Ecuadorian Andes (K. King, 2000). Similar management may be found in the attempts to spread the use of Celtic languages in Ireland, Scotland, Wales, Cornwall, and Brittany. In the case of Cornish, extinct for nearly two centuries, and with only medieval texts available, the decision to revive the language required a major effort to rebuild the vocabulary.

Another example showing how corpus management fosters the status of marginalized languages is the use of institutional power to promote selected languages. For instance, the status of Hebrew was transformed from a sacred and literary language in the Jewish diaspora to the recognized language for everyday use in the Jewish community in Mandatory Palestine, a development which prepared for it for its role as the official language in Israel (Spolsky and Shohamy, 1999). Classical Arabic has been used to enhance religious identity among Muslims; in the Arabic speaking countries, it is usually listed alongside religion in the constitutional definition of the state. After Indian independence, Sanskrit was listed in the Constitution as one of the official languages of India, was proposed as a medium for the transmission of news on the radio, and is now set as the medium of instruction in three universities. The new

draft constitution of Iraq moves Kurdish close to equality with Arabic. An Official Maori Language Act in New Zealand provided official status for what was considered an endangered language. Official status is not enough – Ireland and New Zealand and Valencia were able to increase the numbers who have learned Irish and Maori and Catalan by adding resources for teaching the languages, but they have not so far been able to increase the number of people actually using the language.

The converse of the policy of promoting little-used languages is the deliberate removal or downgrading of languages. The systematic suppression of the use of Tibetan in China or Kurdish in Turkey or the autochthonous languages among the American North, Central, and South American Indians are clear examples. In the same vein, but less dramatic, are the efforts by the former Soviet states (Hogan-Brun and Ramoniene, 2003; Landau and Kellner-Heinkele, 2001) to replace Russian with their titular language. While most of this policy is directed at language use – in government, the press, the media, the educational system – it also includes changes in the language itself. These changes include the purging of Russian forms and vocabulary (added during the period of Stalinist Russification) from the titular language, a search for alternative cultural and historical roots – in the case of Muslim states emphasizing Turkic origins – and the creation of neologisms both to replace Russian borrowings and to modernize the traditional language (Grenoble, 2003). In most of these nations, shifts to roman scripts had begun even before independence, but then became more pervasive. In some of the former satellite Baltic and Eastern European states, the Russian language has been stripped of its dominant position in government and the educational system.

Status Management

The status of a language variety refers to the domains and extent of its use and to its associated rankings in society. More particularly, status management usually refers to the designation of languages as official for use in the public sector and in the educational system. Most scholarly analysis of language planning and policy is concerned with status, although, as Fishman (2000) points out, status and corpus are usually intertwined. The nature of status policy depends substantially on differences in the number and types of languages spoken in a country (Lambert, 1999). Countries with a single dominant language face a different set of policy issues compared with linguistically dyadic or triadic countries – those with two or three relatively equal languages. Similarly, countries that are linguistic-mosaics, that have a large number of significant languages, have different sets of problems from monolingual and dyadic or triadic language countries.

Ideologically Monolingual Countries

Few countries are truly linguistically homogeneous – Iceland is probably the closest (Vikor, 2001) – but many countries in Western Europe, the Americas, and Asia have perceived themselves as being essentially monolingual. In Europe, this is especially striking in the face of persistent multilingualism. The recently published *Encyclopedia of the Languages of Europe* (Price, 2000) listed some 300 historical and currently used languages in Europe. Several of the countries in East Asia, too, essentially see themselves as monolingual although each contains important language minorities. In these cases, there is generally what Fishman (1969) labeled a single Great Tradition, which is associated with a single language. Generally, these countries leave the interest in the selection and standardization of the national language to the field of historical linguistics, but for those with a more recent history, there are studies looking at the first congress proclaiming the language (Fishman, 1993) and of the struggle for standardization. In linguistically homogeneous countries (Fishman, 1966a), the principal focus of language policy has been on corpus management, the cultivation and purification of the national language (for instance, Pedersen, 2003), supplemented in some countries – notably France, Germany, and Japan – by efforts to export the national language abroad (language diffusion policy) (Cooper, 1982).

Within ideologically 'linguistically homogeneous' countries, language policies that relate to linguistic minorities depend in part on the kind of minority involved.

Ethno-Linguistic Regional Minorities Long-standing, geographically concentrated minorities with a recognized history and culture receive the bulk of attention in both governmental and educational language policy, as well as in academic analysis. Examples of such minorities are the Swedish-speaking minority in Finland, the Sami in Finland, Sweden, and Russia, and the Celtic language communities in Ireland, Great Britain, and France. The dominant paradigm in European status policy and in academic analysis is the protection of such linguistic minorities against the absorptive effects of the dominant national language.

A wide variety of country and language specific case studies is now available (Dorian, 1998; Fishman, 2002; Grenoble and Whaley, 1998; Hale, 1991; Krauss, 1991). Most of them exemplify this approach. The use of terms such as 'threatened', 'dying', 'endangered' languages and at the extreme, 'language death',

and 'linguistic genocide' reflect the nature of such analyses. The intended effect of these terms is to characterize the aspirations of ethno-linguistic minorities in terms of group and individual rights. These rights are elaborated by law in many monolingual countries, as well as in covenants and resolutions enacted by international bodies: The European Charter for Minority or Regional Languages, a Framework Convention for the Protection of National Minorities, The Oslo Recommendations regarding the Linguistic Rights of National Minorities, The Hague Recommendation Regarding the Education Rights of National Minorities, and the Universal Declaration of Linguistic Rights (Ferguson, 1968; May, 2001; Nic Shuibhne, 2001). For instance, the term 'other languages' of Europe is a product of an international organization, the European Union. It refers to "all languages apart from the eleven official languages that are ignored in public and official activities of the European Union (Extra and Gorter, 2001: 1)." In practical European Union policy, with its first principle of national sovereignty, the identification of a protected linguistic minority is reserved to the founder states, which have the option to exclude any variety they label as a dialect, as Sweden in 1995 decided that the Charter applied to Sami, Tormedal Finnish, Finnish, Romani Chib, and Yiddish, but not Skanian with 1.5 million speakers, and France prefers not to recognize Occitan.

The effect of official designation of a minority language, whether within a country or internationally, can be of substantial benefit to the group, expanding its claim to educational and governmental support. Consequently, there is constant pressure to expand the list, drawing the line further down the continuum from language to dialect or giving legal identity to different types of languages. For instance, the deaf community has sought recognition of sign languages as separate minority languages, but the European Union continues to resist this. Efforts have also been made in the United States to declare Black English a minority language, and thus subject to special protection. There has been a movement to imbed the concept of language rights in a larger framework, the promotion of multilingualism for the general population, autochthonous or immigrant (Skutnabb-Kangas, 2000).

Policies toward linguistic minorities differ according to their relative size, their degree of geographic concentration, their historical roots, their extra-country linkages, the strength of their ethnic identification, and the political activism of their leadership (Paulston, 1994). The features of official language policy that vary according to these characteristics are: (a) a language's role in the education system, in particular the class and school levels in which it is used, and whether it is taught as a subject or used as a medium of instruction; (b) its role in governmental affairs – the legislature, judiciary, administrative services, the military; (c) its role in the media, particularly that portion controlled by government; (d) the possibility of using it in access to governmental and commercial institutions, and (e) its use in the workplace.

In academic analyses of minority language policy, a number of constructs have been proposed to arrange language minorities along continua of relative vitality. A widely used scale is the Graded Intergenerational Disruption Scale (Fishman, 1991) based upon a language's presence in governmental affairs, education, adult use, and intergenerational transfer. The scale also purports to advise linguistic minorities on how to advance their status and how to promote the use of the language. It ranges from the most threatened eighth stage, where any effort needs to start with 're-assembling' the language from 'vestigial users . . . socially isolated old folks' and teaching it to adults, to the highest stage, where the language is used to some extent in 'higher level educational, occupational, governmental, and media efforts' but lacks the safety of political independence. In between, there are another half dozen levels, the most significant of which are probably the sixth ('intergenerational informal oracy'), the fifth ('institutional unsupported literacy') and the fourth ('use in official lower education'). In Fishman (2001), where various scholars are asked to comment on the scale, several raise questions about the ordering of the scale: for instance, there are many cases where institutionalized literacy teaching (commonly of a religiously sanctified language) continues even without much everyday oral use.

Fishman's scale was developed to account for the process that he labeled Reversing Language Shift, an attempt by supporters of a language to re-establish or establish its status. Also named 'language revival', the process of re-establishing natural intergenerational transmission (language revitalization) or vernacular use of a literary language (revernacularization) is most clearly exemplified by Hebrew.

The success of the Celtic revivals in various countries has depended on the extent to which they are backed by political power, as in Ireland where the Celtic language, Gaelic, has become a symbol of nationhood, or in Wales, where a regional government has championed its use. However, even in countries and regions where there is strong governmental backing, only a minority of the population actually speaks the Celtic language. Other cases of revival involve territorial linguistic minorities.

Territorial Linguistic Minorities Territorial linguistic minorities also differ in the extent to which their speakers seek full political autonomy, as do the Tamils of Sri Lanka, and some of the Basques in Spain. For most groups, however, the goal is limited to the use of the minority language in governmental affairs and at various levels of the education system. For instance, in Spain in three constitutionally mandated autonomous regions, Basque, Catalan, and Galician languages are not only taught in schools, but public use of the language is actively promoted, and, because their speakers occupy their own political units within Spain, they can determine their own official language policy within their territory (Turell, 2001). By way of contrast, in France, the Basque-speaking sections bordering on Spain are not officially recognized as separate language groups, they do not comprise a separate political unit, and they cannot determine linguistic policies. In France, the promotion of the Basque language is left to voluntary initiatives. In a similar vein, the various Celtic languages represent different kinds of territorially specific language minorities with varying claims on governmental power – one result of autonomy for Wales and Scotland has been to boost the claims of Welsh and Scottish Gaelic. In New Zealand, the campaign for Maori language regeneration accompanied a series of legal claims before a Tribunal set up to remedy failures to carry out provisions of the 1840 Treaty of Waitangi, most of which sought financial reparations for lost land and hunting and fishing rights (Spolsky, 2003).

There are many other territorially-concentrated linguistic minorities elsewhere in Europe, such as the Frisians in the Netherlands who receive special treatment in support for their bilingual education but not in their dealings with government (Gorter, 2001). Special accommodation is also made for territorial linguistic enclaves whose residents are speakers of languages of neighboring countries. For instance the Swedes in the southwestern corner of Finland (Vikor, 2000), and the Germans in the contiguous border regions of Belgium (Aunger, 1993), Italy, and France are examples of transborder linguistic minorities, the former deriving recognition from historic political union and continuing territoriality.

There are a few long-established linguistic minorities that are not geographically concentrated and that typically receive less policy attention. Of these, the most notable are the Roma or Romani, who are scattered and peripheralized. In its negotiations with new candidates for membership, the European Union generally exerted considerable pressure to have these languages supported, following a principle of 'do what I say and not what I do' and not giving them the privilege accorded to foundation members of choosing which varieties to support. While the main goal of Chinese language policy has to do with developing the common language (Putonghua) and in simplifying the characters used for writing Chinese, the cultivation and preservation of minority languages has now been added as a goal (White, 1997). After most speakers of Yiddish in Europe have been killed or have emigrated, some European countries now recognize Yiddish as a minority language (Hult, 2004).

Paulston (2004) has proposed what she calls "extrinsic linguistic (or ethnic) minorities," groups such as the Russians in Baltic Republics who went from being majorities to minorities by legal measures or the moving of borders or grants of independence, but who continue to show strong language loyalty.

Aboriginals Like other territorially concentrated linguistic minorities within homogeneous states, culturally distinct autochthonous groups receive a great deal of attention both in language policy and in academic analyses. Often the languages of such groups are in a wide variety of stages of development. Hence, a primary focus of management is on alphabetization and the promotion of literacy and oracy. In most cases, the drive for language rights among aboriginal groups is tied to cultural revival and reinforcement. Linguistic groups whose members are still active speakers of their languages and who are territorially concentrated, such as the Samis in the Nordic countries (Jernsletten, 1993) and Russia, or the Quichua (K. King, 2000) in the Andean highlands of Peru, Bolivia, and Ecuador have greater success in achieving special treatment in language policy. More dispersed aboriginal groups such as the American Indians and the aboriginal tribes in Australia who are dispersed through a hundred different regions (Lo Bianco and Rhydwen, 2001), have an even greater difficulty in language maintenance – although the Navajo have had some success (McCarty, 2002; Spolsky, 2002). An exception are the Maoris in New Zealand who have had great success in cultural and linguistic revival through concentrated political agitation and through the use of Maori in Te Kohanga Reo, the preschool 'language nest' programs and the subsequent development of immersion education in elementary schools (Benton and Benton, 2001). Autochthonous minorities, although commonly suffering from political and social and economic discrimination, at least can claim that they were there first. Autochthonous languages are obviously especially endangered, for they lack other territories where they are spoken.

Immigrants Language policies are much less accommodating to the needs of immigrant groups. In fact, almost all of the international covenants supporting the rights of linguistic minorities apply only to citizens, arguing that immigrants chose to live in the country and so can reasonably be expected to make an effort to learn its language. In addition, the claim that the preservation of immigrant languages is required for maintenance of language diversity is a weak one, for their languages are usually spoken in the country from which they come. However, recently this distinction has been blurred (Hornberger, 1998). In the early years immediately after World War II during the first major flow of 'guest workers' into Europe, they were expected to go back to their home countries after a brief sojourn. Moreover, at that time, any service of their linguistic needs in education was supposed to be provided by their home countries. In addition, immigrant groups tended to be widely dispersed in cities and did not constitute a separate territorial unit. Over more recent decades, their numbers have grown immensely, particularly with Eastern Europeans migrating into Western Europe, citizens of former colonies moving to the metropolitan homeland, and a major wave of migrants from Islamic countries. As their numbers have grown, they have not tended to form separate territorial groups, although their concentration in urban areas, their numbers, and their growing political influence have come to require special educational and governmental accommodation. These may include the provision of instruction in the home language in primary schools, the translation of government documents and court proceedings into the home language, and, in some countries, support for instruction of new immigrants in the national language of the country. In spite of expressions of support for immigrants and their human rights, there has been a tendency to require proficiency in the official language for citizenship and in some cases for immigration.

The United States provides a clear example of this transformation. Over two centuries, massive waves of immigrants have been absorbed. Historically, they tended to be widely dispersed into a number of cities, where little islands would be created. Each group, however, was expected in time to merge into the general population, including the learning of English (Fishman, 1966b). After a period when immigration was restricted by legal quotas, the number of immigrants has increased rapidly. As a result, there are now 3 million children in the United States who speak at home a language other than English. They are referred to as Limited English Proficiency (LEP) children. Three-fourths of the LEP students are Hispanic, and instead of dispersing throughout the country they have become a major territorial language minority in Florida and the American southwest and West, particularly California. One result was the institution of language rights accorded territorial linguistic minorities elsewhere, including a highly institutionalized system of bilingual education in primary schools, and representation of Spanish in public life and the media (Roca, 2000). This development has given rise in some states to reverse pressure to enact legislation banning bilingual education and making English the only official language (Baron, 1990).

Dyadic or Triadic Societies

Countries that have two or three major recognized languages such as Canada, Belgium, Switzerland, Sri Lanka, and Cyprus, each with its own territory, have problems of language policy different from those facing ideologically homogeneous countries. In such countries, language management issues tend to pervade large sectors of the educational system and public life. As in linguistically homogeneous counties, some provision may be made for lesser language minorities, but the fabric of the state itself tends to be linguistically consociational involving only the primary languages. The preferred solution to any conflict is territorial: governmental and educational institutions are organized separately in the different language areas, and political power is carefully balanced between the linguistic units. An extreme example is Belgium, a country historically formed by uniting monolingual territories. After four governmental crises based on language issues between 1979 and 1990, the country was partitioned into different language regions: (a) areas that are exclusively monolingual in Dutch or French, (b) areas such as Brussels that are officially bilingual, and (c) areas that are monolingual but provide some minority language rights (Deprez and Du Plessis, 2000). Switzerland has a longer-established form of consociational linguistic territoriality, but restricts its implementation primarily to educational and governmental affairs; each of the 27 cantons is autonomous in language choice. Canada, too, was formed out of previous, distinct French and English territories. To maintain unity, it is formally bilingual, but French-speaking Quebec periodically attempts to gain independence from the other, primarily Anglophone-speaking, provinces of Canada. A series of referenda for Quebec's independence has not gained a majority of votes in Quebec, defeated by negative votes from a combination of Anglophones, aborigines, and immigrant communities. However, in Quebec province itself, the use of French in all governmental affairs, education, and public displays is mandated (Bourhis,

2001). In Anglophone Canada, an innovative policy was introduced whose intent was to disarm the Quebec separatist drive. Schools for non-francophones require their students to enroll in immersion classes to make them proficient in French. This widely watched program has been only modestly successful.

In some countries, the relationship between the ethno-linguistic groups is so contentious that the country breaks apart, as in the former Czechoslovakia and Yugoslavia (Bugarski, 2001). In post-independence Pakistan, two linguistically different sectors were separated by a thousand miles – a Bengali-speaking Eastern half and an Urdu-, Punjabi-, and Sindhi-speaking western half (Rahman, 2002). After a bitter war, the eastern sector became a separate country, Bangladesh. A two millennia-old conflict between Tamils and Sinhalese in Sri Lanka is in danger of partitioning the island into two countries, as is the conflict between the Greece- and Turkey-oriented halves of Cyprus. Sometimes in binary societies, one language group dominates the other as in the Sudan where the Arabic-speaking North dominates the lower multilingual, tribal-based south, or in Israel, officially bilingual in Hebrew and Arabic, where Arabic (though benefiting from more use in education than in many nominally Arab countries (Amara and Mar'i, 2002) is clearly dominated by Hebrew.

In dyadic nation states, then, the key management problem usually remains the resolution of competing demands for status between two languages with strong claims.

Mosaic Societies

Most countries are neither homogeneous nor dyadic nor triadic in composition. Indeed, the majority of countries in the world are made up of five or more important ethno-linguistic and territorially discrete segments. The problems of language policy, both corpus and status, in mosaic countries such as India, Indonesia, the Philippines, and most of the countries of Africa are immense and complex. Here, status and corpus are inexorably linked: a language's claim to official recognition is clearly bounded by its state of cultivation, for it is difficult to use an unwritten language in schools or an unmodernized language to teach science. In many of these countries, the overwhelming primary concern is corpus management, in particular the development of a written form of the languages, the promotion of literacy among the public, modernization by developing new terminology, the staging and duration of language instruction at the various levels of the educational system, and the preparation of teaching materials and teachers. Moreover, the solutions to status policy issues that are available in ideologically monolingual, dyadic, or triadic countries do not apply where there are many languages.

In mosaic societies, even the number of languages spoken in a country is often uncertain. Various counts have enumerated between 1000 and 2000 languages in Africa. In Nigeria alone, a variety of linguistic censuses have found 200 to 400 languages. At last count, there are 535 languages in India. In the late 19th century, Grierson counted a thousand. In all of these countries, the number of language varies immensely in part because the dividing line between languages and dialects is indistinct and political rather than linguistic.

Those who wish to develop language policy in such countries face a number of special challenges. In many of them, a single over-arching language was introduced by the former colonial power and is still used by a small elite. There were two major approaches to colonial language policy. France and Portugal (like Spain in Latin America) were consistently ruthless in requiring the metropolitan language for all government and for any education they supported. After its experience in India, Britain in other parts of the world followed what might be called a modified Oriental policy, providing initial education (at least the first two or three years and sometimes up to secondary school) in reasonably populous indigenous vernacular languages, at least those with a writing system. German and Belgian colonial policy similarly allowed a small place for vernacular languages. After the primary level, both approaches then accepted the centrality of the metropolitan language, but the British did encourage some continued cultivation of some indigenous varieties.

In the optimistic days after World War II, post-colonial political pressure was to dethrone the colonial language and nativize the choices of national languages. A number of African and Asian former colonies started to indigenize their schools. However, the initial pressure for abolishing the colonial tradition has had to be balanced against the tendency among indigenous elites in many former colonial countries to distinguish themselves by their command of the colonial language (Myers-Scotton, 1993), and increasing proportions of the population see the command of that language as the path to upward mobility. Moreover, the exclusive choice of native languages sacrifices links to modernity and international communication. As a result, the use of ex-colonial languages lingers and may be growing. For instance, while the Indian constitution prescribes that English was to be abandoned as a national language within ten years, it still remains one of the official languages. Moreover, Indians of all social

classes see the mastery of English as the avenue for upward mobility, and enrollment in English-medium private schools is growing (Dua, 1996). Similarly, in most former Francophone states in Africa, French remained the official language after independence, to be threatened most recently by globalizing English rather than by local national languages (Chumbow and Bobda, 2000).

The process of nativization, with its shift to indigenous languages, is handicapped by the number of those languages and their regional or tribal identification, with all of the status implications resulting from selection of one or a few languages and so favoring its speakers over others. Solutions adopted in a variety of countries include the creation of a fresh lingua franca, usually adopting a local dialect, often one close to the capital city, or adopting a regional language. The use of the new lingua franca is then promoted for use in the education system, in government, and in the media. One of the most striking examples is Bahasa Indonesia, developed out of Malay and now the national language (Dardjowidjojo, 1998). Malay was also the basis of Bahasa Malaysia, and the slight variant Bahasa Melayu developed in Malaysia and Brunei, but there are new pressures for English to be used there (Omar, 1998). Other cases are Tok Pisin, in Papua/New Guinea, Filipino, a variant of Tagalog, in the Philippines, and the adoption of Swahili in Tanzania and East Africa. It should be noted that in Malaysia, there has been a decision to move to English-medium instruction at all educational levels (Gill, 2002).

Many mosaic countries have chosen a language policy model which reflects one or another stage in the history of language policy in the former Soviet Union (E. Lewis, 1972). In the early Soviet period, the languages of the 15 principal language regions were declared to be of equal status. Each was declared the official language and taught in the schools in its own region. Every child had the right to be educated in his or her own language. Russian was to be *primus inter pares*. The decision to encourage and cultivate the vernaculars was based on the principle that it would be the fastest way to develop communism among illiterate peoples, and Grenoble (2003) notes that this policy did result in the rapid development of literacy. Under Stalin, with the pressure for central control, the status of the regional languages was downgraded and the spread of Russian was promoted.

India initially adopted the Soviet model. At Independence, the boundaries of the states were redrawn from the multilingual units they had been under British rule to more or less monolingual units, taking into account the major literary languages, as the political parties in the independence movement had urged. In the years immediately after independence, there was a great deal of concern in India about the balkanizing effect of this decision. To combat what were called 'fissiparous tendencies' Hindi – a Sanskritized form of Hindustani – was chosen to be the bridging national language. However, the states in southern India, whose languages belong to an entirely different family, strongly objected. As happens in many mosaic societies, the resulting compromise piled on languages in the educational system. The medium of instruction in the primary school was to be the local language, with various other languages added in secondary and higher education, serving as either media of instruction or as subjects of study. India's compromise was called the Three Language Formula – in primary school the local language would be used; in secondary school Hindi, English, and the regional language would be taught. In the Hindi area in the north, another regional or European language was to be substituted. As yet this policy has not been rigorously applied and, *de facto*, the local languages still seem dominant with English serving as the bridge language. While such compromises mitigate political difficulties in mosaic countries, the problem of governmental communication remains, particularly which languages can be used in governmental affairs. This usually requires the adoption of one or a few working languages, or allowing the use of many languages but providing a mechanism for interpretation and translation (Itagi and Singh, 2002).

The People's Republic of China essentially continued a 2000-year old tradition for Chinese languages by continuing the ideology that they were all dialects, united by their single writing system (Zhou, 2004). Language management then became a matter of finding a way to simplify the characters, supplement them with a more or less phonetic alphabet, and encourage a shift to Putonghua, the variety of Mandarin based on the Beijing dialect. For the non-Chinese languages, the initial policy was based on the Soviet model, with the development of literacy in and recognition of a manageable number of varieties, originally (as in the Soviet approach) selecting one dialect as the basis of standardization. There was, however, no effort to force them to accept the Chinese writing system, but rather acceptance of various traditional scripts or use of modified Roman or Cyrillic alphabets. At one period, there was a strong effort to assimilate these groups, too, linguistically as well as culturally, but more recently, there is an acceptance of bilingual solutions for the larger languages (Zhou, 2003).

Most African nations are afflicted with the effect of the lack of congruence between imperially established and tribal or linguistic boundaries. They generally

include many languages, many of which are spoken by large numbers in bordering states. In former French colonies, the position of French as language of government and advanced education is well established (Salhi, 2002), in spite of efforts in North Africa to establish the status of Arabic (Daoud, 2001). Portuguese, too, remains dominant in former Portuguese colonies, though in some countries Creoles are developing and becoming important (Vilela, 2002). In West Africa, Bamgbose (2000) complains, there are national language policies that do not reflect an understanding of local linguistic practices and that are seldom seriously implemented. Illiteracy rates are high; colonial language policies mainly remain in effect. In Botswana, Nyati-Ramahobo (2000) reports, the indigenous languages other than Setswana have been ignored or discouraged, and English is favored over it in government and education. Summarizing the current status in Africa, Batibo (2004) notes that only 2 countries (Egypt and Libya) have adopted indigenous languages as their official medium, 8 use an indigenous language alongside an ex-colonial language, 27 use an ex-colonial language with some symbolic secondary use of an indigenous language, and 18 have ex-colonial languages as the only official national language. In other words, 80% have failed in any efforts to establish indigenous languages as official languages. Two countries that use indigenous languages in the school system, Botswana and Tanzania, require its use by all students, whatever their mother tongue. Many are hopeful that the recognition of a number of indigenous languages in the South African constitution alongside English and Afrikaans will lead to multilingual policies (Kamwangamalu, 2000; Mesthrie, 2002), but studies are suggesting how slow the process is (Heugh, 2003).

This same problem of mosaic societies is not limited to single countries, but faces international organizations with sovereign states as members who must communicate in multilingual contexts. The Council of Europe, for instance, now has 45 member states. It has adopted French and English as its official languages of communication. The United Nations publishes its daily journal in English and French, but has six 'working languages' in which official statements may be made: Arabic, Chinese, English, French, Russian, and Spanish. If a delegation wants to communicate in another language, it must provide translators and interpreters. The European Union provides for translation and interpretation among the languages of all its members, requiring about 200 simultaneous translators for a single session if all possible language pairs are to be covered.

The issue of language policy in international organizations with their presumption of equality among the languages of member states illustrates the more general problem of the tension between status considerations in language choice and the need to make communication in multilingual contexts effective (Ammon, 2002; de Swaan, 1999; Van Els, 2001). The *de facto* primacy of English as the language of communication is not without its critics (Phillipson, 2003). Elsewhere, when the need for international communication is paramount, the trend is to use English as the common language. For instance, 85% of the citations in the world's scientific literature are published in English (Garfield and Alfred, 1990). While the multilingual capacity of the computer and the Internet has belied the prediction that only English would be used, the pressure for English appears to continue unabated. The growing predominance of English in international communication, of course, is a major handicap to speakers of other languages, and there are numerous attempts such as 'English as a Lingua Franca' or 'World English' to modify the language to make it more accessible to non-native speakers. The perennial attempts to foster the use of Esperanto serve the same purpose.

Foreign Language Teaching Policies

There is some overlap in language teaching policy between domestic and foreign languages. For instance, French is both a domestic and foreign language in Canada, as are French and English in many ex-colonial countries. However, in the main, foreign language policies are usually quite distinct from, and less developed and conflicted than, policies with respect to national language(s) and those of intra-country minorities. They also tend to be given less attention in scholarly analyses of language policy. In addition, such policies tend to be piecemeal rather than coordinated. Only a few overall national foreign language teaching policies have been adopted. The national plan for The Netherlands (van Els, 1992) is one of the few that were based upon surveys of adult use and national need. Australia's national policy statement (Lo Bianco, 1987) included policies with respect to indigenous peoples and immigrants as well as foreign languages. Comprehensive national policy in England was until recently either expressed as part of official curricula for all instruction, or is addressed by non-governmental organizations (Moys, 1998), including a new policy document developed from the 2000 Nuffield Report (Department for Education and

Skills, 2004). Foreign language policy normally relates only to the educational system, although France and Egypt try to limit the use of English outside the educational system. Within the education system, there are a number of common issues that foreign language policy must face (Bergentoft, 1994).

One basic decision concerns the proportionate role of foreign language instruction in the curriculum. In most mosaic countries, the promotion of multilingualism in intra-country languages and perhaps the colonial language leaves little time for foreign languages. The study of foreign languages is most fully developed in Western Europe, where statutory mandates usually require the study of one, and in most countries, two, foreign languages. It appears that the reason for a two-foreign-language policy is to ensure that languages other than English, which is almost always the first choice, are included. Language study may take up a substantial proportion of curricular time. In Sweden, for instance, language study may absorb 15% of total curricular time. In Luxembourg, where French, German, English, in addition to Luxemburgish, are required, the proportion of time taken up in language study is much higher.

Time spent on foreign language study is generally less in the English-speaking countries (Moys, 1998). In the United States, although some state governments which have authority over education do mandate the teaching of foreign languages, the decision on how much foreign language should be offered is usually left to individual districts and schools. All 50 states include the study of foreign languages in their secondary school curricula, although no state requires the study of foreign language in secondary school as a graduation requirement for all students, and only ten states require language study for college-bound students. Unlike other countries, in the United States students may start their language study in higher educational institutions. In 2002, there were 1.4 million students enrolled in foreign language classes in 780 colleges and universities. However, unlike other countries where students enroll in foreign language study in primary school and continue throughout secondary school, enrollments in the United States foreign language classes tend to start in secondary school or college, and drop on the average by half from one language course level to the next. In many countries there is an increasing tendency to start language study earlier and earlier in primary schools, but the practice is still uncommon in the United States. In England, where a decision was made to drop the requirement for foreign language study after the age of fourteen in comprehensive schools, and in the United States, where budgetary pressures became intense, the number of foreign language courses dropped precipitously.

Foreign language teaching policy specifies which languages are to be studied and in what order of priority. This choice is determined by government fiat in some countries. In many countries, however, school and student choices are primary. In England, and formerly in the United States, the traditional order of language selection for modern languages was French and then German. In the United States, Spanish has become the overall favorite, with French and German in steep decline. French, a language spoken in a country a short journey away, remains the favorite in England. In the other countries of Western Europe, the language chosen after English is likely to be German, followed by French and Spanish. In almost all non-English speaking countries that require foreign language study, the first language to be studied is English, selected by eighty per cent or more of the students, often starting in primary school (Bergentoft, 1994). In the United States, federal governmental support, provided during the Cold War for the teaching of Russian, now promotes the study of the languages of Asia and the Middle East at the higher education level; this support has been boosted since 9/11. Except for instruction specifically aimed at immigrants, Asian languages are seldom taught in countries outside of their home regions.

While some countries specify the method of teaching in language classrooms, in the main, the choice of style of classroom instruction is left to teachers, school districts, and textbook publishers. Indeed, the general trend is away from centralized control of language education to more localized and individual teacher decisions. There are, however, some general trends in the style of language teaching that are taking place in most countries. Particularly in Europe there has been a tendency toward the adoption of what is called communicative competence-oriented language instruction and the primacy of oracy over reading and writing skills. Moreover, the Council of Europe has been instrumental in bringing about a modernization and uniformity in language teaching in many countries. In the early 1990s, what was referred to as the Threshold Level (van Ek, 1975) was introduced by the Council of Europe. It provided specific communicative competence goals that students were expected to achieve. The Threshold Level has been adopted throughout Europe for the teaching of 20 languages, and more advanced levels have been described (Ek and Trim, 1991, 2001). The Council also provided to its members a widely adopted series of guidelines for everything from teacher training, elementary school language instruction, and language education

for vocational students. The European Union supported research throughout Europe on improvement of foreign language teaching and provided advice on general language instructional strategies to all of its member states (European Commission, 1997).

Much of the control of the nature of foreign language instruction lies with the adoption of uniform strategies for assessment. In this regard, once again the international organizations in Europe have been helpful. The Council of Europe developed a set of language assessment standards, the Common European Framework (Council of Europe, 2001) intended to promote a degree of uniformity among its members, with a goal of facilitating the growing practice of student exchanges (Scharer and North, 1992). These standards have been widely adopted throughout Europe and are influential elsewhere. In the United States the most important, indeed the only, national attempt to make uniform policy for foreign language instruction is the development of a set of standards for a substantial number of languages. Developed by a teachers' organization, the American Council on the Teaching of Foreign Languages, the ACTFL guidelines (American Council on the Teaching of Foreign Languages, 1986) has had a major effect on the modernizing of foreign language instruction throughout the United States.

In China, after the end of the Cultural Revolution and even more with the access to the World Trade Organization, there has been a centrally mandated increase in foreign language teaching, with an emphasis now on English (as opposed to an earlier emphasis on Russian) but with a wide choice of other languages. Methodology, too, is being revised, with a new concern for oral language and for humanistic approaches.

Conclusion

In summary, both the development and analysis of language policy have grown immensely in the past several decades. Earlier interest in corpus management has now been overshadowed by a surge of interest in status policy, particularly as it relates to the rights of territorial, regional, and aboriginal minorities. There has also been an increase of interest in language acquisition policy, but it still receives less attention and is almost entirely unrelated to the rest of language policy. However, anyone following the topic in the world's press can usually find two or three stories about language policy a day, and scholarly activity is burgeoning to keep up. It seems safe to predict that the study of language policy in general will continue to develop rapidly.

See also: Endangered Languages; Identity and Language; Language Ideology; Language Loyalty; Language Maintenance and Shift; Language Policy in Multilingual Educational Contexts; Linguistic Decolonization; Minorities and Language; Nationalism and Linguistics; Norms and Correctness; Politics and Language: Overview; Power and Pragmatics; Reversing Language Shift; Spelling Reform.

Bibliography

Amara M H & Mar'i A A-R (2002). *Language education policy: the Arab minority in Israel.* Dordrecht, The Netherlands: Kluwer Academic Publishing.

American Council on the Teaching of Foreign Languages (1986). *ACTFL proficiency guidelines.* New York: American Council on the Teaching of Foreign Languages.

Ammon U (2002). *Present and future language conflict as a consequence of the integration and expansion of the European Union (EU).* Paper presented at the XXXVI Congresso Internationale di Studi della Societa di Linguistica Italiana, Bergamo.

Annamalai E (1989). 'The linguistic and social dimensions of purism.' In Jernudd B & Shapiro M J (eds.). 225–231.

Aunger E A (1993). 'Regional, national and official languages in Belgium.' *International Journal of the Sociology of Language 104,* 31–48.

Bamgbose A (2000). 'Language planning in West Africa.' *International Journal of the Sociology of Language 141,* 101–118.

Baron D E (1990). *The English only question.* New Haven: Yale University Press.

Batibo H (2004). *Marginalisation through educational medium: the case of the linguistically disadvantaged groups in Botswana and Tanzania.* (LAUD Paper 605). Essen, Germany: University of Duisburg-Essen.

Benton R A & Benton N (2001). 'RLS in Aotearoa/New Zealand 1989–1999.' In Fishman J A (ed.). 422–449.

Bergentoft R (1994). 'Foreign language instruction: a comparative perspective.' In Lambert R D (ed.) *Language planning around the world: contexts and systemic change.* Washington, D.C.: National Foreign Language Center. 17–46.

Bourhis R Y (2001). 'Reversing language shift in Quebec.' In Fishman J A (ed.). 101–141.

Bugarski R (2001). 'Language, nationalism, and war in Yugoslavia.' *International Journal of the Sociology of Language 151,* 69–87.

Chumbow B S & Bobda A S (2000). 'French in West Africa: a sociolinguistic perspective.' *International Journal of the Sociology of Language 141,* 39–60.

Cooper R L (1989). *Language planning and social change.* Cambridge: Cambridge University Press.

Cooper R L (ed.) (1982). *Language spread: studies in diffusion and social change.* Bloomington, IN: Indiana University Press.

Coulmas F (1991). 'The future of Chinese characters.' In Cooper R L & Spolsky B (eds.) *The influence of language on culture and thought: essays in honor of Joshua*

A. Fishman's sixty-fifth birthday. Berlin: Mouton de Gruyter. 227–243.

Council of Europe (2001). *Common European framework of reference for languages: learning, teaching, assessment.* Cambridge: Cambridge University Press.

Cummins J (ed.) (1983). *Heritage language education: proceedings of a conference organized by the Multilingualism Directorate of the Department of the Secretary of State, Saskatoon.* Canada: Multiculturalism Canada.

Daoud M (2001). 'The language situation in Tunisia.' *Current Issues in Language Planning 1(1),* 1–52.

Dardjowidjojo S (1998). 'Strategies for a successful national language policy: the Indonesian case.' *International Journal of the Sociology of Language 130,* 35–47.

DasGupta J & Ferguson C A (1977). 'Problems of language planning.' In Rubin J, Jernudd B, Das Gupta J, Fishman J A & Ferguson C A (eds.) *Language planning processes.* The Hague: Mouton. 3–7.

de Swaan A (1999). *The language constellation of the European Union.* Paper presented at the International status and use of national languages in Europe: contributions to a European language policy, Brussels.

Department for Education Skills (2004). *Languages for all: languages for life: a strategy for England.* London: Department for Education and Skills.

Deprez K & DuPlessis T (eds.) (2000). *Multilingualism and government: Belgium, Luxembourg, Switzerland, Former Yugoslavia, and South Africa.* Pretoria, South Africa: Van Schaik Publishers.

Dominguez F & López N (1995). *Sociolinguistic and language planning organizations.* Amsterdam and Philadelphia: John Benjamins Publishing Company.

Dorian N (1998). 'Western language ideologies and small-language prospects.' In Grenoble L A & Whaley L J (eds.) *Endangered languages: current issues and future prospects.* Cambridge: Cambridge University Press. 3–21.

Dua H R (1996). 'The spread of English in India: politics of language conflict and language power.' In Fishman J A, Rubal-Lopez A & Conrad A W (eds.) *Post-Imperial English.* Berlin & New York: Mouton de Gruyter. 557–588.

Ek J A van & Trim J L M (1991). *Waystage 1990.* Cambridge: Cambridge University Press.

Ek J A van & Trim J L M (2001). *Vantage.* Cambridge: Cambridge University Press.

European Commission (1997). *Learning of modern languages at school in the European Union.* Luxembourg: Office for Official Publications of the European Communities.

Extra G & Gorter D (eds.) (2001). *The other languages of Europe: demographic, sociolinguistics and educational perspectives.* Clevedon, England/Buffalo, NY: Multilingual Matters Ltd.

Ferguson C A (1964). 'Baby talk in six languages.' *American Anthropologist 66(6.2),* 103–114.

Ferguson C A (1968). 'St. Stefan of Perm and applied linguistics.' In Fishman J A, Ferguson C A & Das Gupta J (eds.) *Language problems of developing nations.* New York: Wiley. 253–265.

Fishman J A (1966a). 'Some contrasts between linguistically homogeneous and linguistically heterogeneous polities.' *Sociological Inquiry 36,* 146–158.

Fishman J A (ed.) (1966b). *Language loyalty in the United States: the maintenance and perpetuation of non-English mother tongues by American ethnic and religious groups.* The Hague: Mouton.

Fishman J A (1969). 'National languages and languages of wider communication in the developing nations.' *Anthropological Linguistics 11,* 111–135.

Fishman J A (ed.) (1977). *Advances in the creation and revision of writing systems.* The Hague/Paris: Mouton.

Fishman J A (1991). *Reversing language shift: theoretical and empirical foundations of assistance to threatened languages.* Clevedon, England: Multilingual Matters Ltd.

Fishman J A (ed.) (1993). *The earliest stage of language planning: the "First Congress" Phenomenon.* Berlin: Mouton de Gruyter.

Fishman J A (2000). 'The status agenda in corpus planning. In Lambert R D & Shohamy E (eds.) *Language policy and pedagogy: essays in honor of A. Ronald Walton.* Amsterdam: John Benjamins Publishing Company. 43–52.

Fishman J A (ed.) (2001). *Can threatened languages be saved? Reversing language shift, revisited: a 21st century perspective.* Clevedon, Avon: Multilingual Matters Ltd.

Fishman J A (2002). Endangered minority languages: prospects for sociolinguistic research. *MOST Journal of Multicultural Studies 4(2).*

Garfield E & Alfred W-D (1990). 'Language use in international research: a citation analysis.' In Lambert R D (ed.) *Foreign language in the workplace.* Philadelphia: The American Academy of Political and Social Science.

Gill S K (2002). *International communication: English language challenges for Malaysia.* Serdang, Malaysia: Universiti Putra Malaysia Press.

Gorter D (2001). 'A Frisian update of reversing language shift.' In Fishman J A (ed.). 215–234.

Grenoble L A & Whaley L J (eds.) (1998). *Endangered languages: current issues and future prospects.* Cambridge: Cambridge University Press.

Grenoble L A (2003). *Soviet language policy.* Dordrecht: Kluwer Academic Publishers.

Guitarte G L & Quintero R T (1974). 'Linguistic correctness and the role of the academies in Latin America.' In Fishman J A (ed.) *Advances in language planning.* The Hague: Mouton. 315–368.

Hale K (1991). 'On endangered languages and the safeguarding of diversity.' *Language 68(1),* 1–3.

Harrison M & Gilbert S (eds.) (2000). *Obscenity and pornography decisions of the United States Supreme Court.* Carlsbad, CA: Excellent Books.

Haugen E (1959). 'Planning for a standard language in Norway.' *Anthropological Linguistics 1(3),* 8–21.

Heugh K (2003). *Language policy and democracy in South Africa: the prospects of equality within rights-based policy and planning.* Ph.D. thesis, University of Stockholm.

Hogan-Brun G & Ramoniene M (2003). 'Emerging language and education policies in Lithuania.' *Language Policy 2(1),* 27–45.

Hornberger N H (1998). 'Language policy, language education, language rights: indigenous, immigrant, and international perspectives.' *Language in Society 27(4),* 439–458.

Hornberger N H (ed.) (1996). *Indigenous literacies in the Americas: language planning from the bottom up*. Berlin: Mouton.

Hult F M (2004). 'Planning for multilingualism and minority language rights in Sweden.' *Language Policy 3(2).*

Itagi N H & Singh S K (eds.) (2002). *Linguistic landscaping in India with particular reference to the new states: proceedings of a seminar.* Mysore, India: Central Institute of Indian Languages and Mahatma Gandhi International Hindi University.

Jernsletten N (1993). 'Sami language communities and the conflict between Sami and Norwegian.' In Jahr E H (ed.) *Language conflict and language planning*. Berlin: Mouton de Gruyter. 115–132.

Jernudd B & Shapiro M J (eds.) (1989). *The politics of language purism*. Berlin/New York: Mouton de Gruyter.

Jernudd B (ed.) (2001). *Language management and language problems: special issue of Journal of Asian Pacific Communications 11:1* (Vol. 2). Amsterdam: John Benjamins.

Kamwangamalu N M (2000). 'The new South Africa, language, and ethnicity: a prognosis.' *International Journal of the Sociology of Language 144,* 137–138.

Kaplan R B & Baldauf R B (2003). *Language and language-in-education planning in the Pacific Basin*. Dordrecht: Kluwer Academic Publishers.

Karyolemou M (2002). 'When language policies change without changing: the University of Cyprus.' *Language Policy 1(3),* 213–236.

King K A (2000). *Language revitalization process and prospects: Quichua in the Ecuadorian Andes*. Clevedon, Avon: Multilingual Matters Ltd.

King R D (2001). 'The poisonous potency of script: Hindi and Urdu.' *International Journal of the Sociology of Language 150,* 43–59.

Kloss H (1969). *Research possibilities on group bilingualism: a report*. Quebec: International Center for Research on Bilingualism.

Krauss, M.(1991). *Endangered languages*. Paper presented at the Linguistic Society of America Annual meeting.

Kroskrity P V (1998). 'Arizona Tewa Kiva speech as a manifestation of a dominant language ideology.' In Schieffelin B B, Woolard K A & Kroskrity P V (eds.) *Language ideologies: practice and theory.* New York/ Oxford: Oxford University Press. 103–122.

Lambert R D (1999). 'A scaffolding for language policy.' *International Journal of the Sociology of Language 137,* 3–25.

Landau J & Kellner-Heinkele B (2001). *Politics of language in the ex-Soviet Muslim states: Azerbaijan, Usbekistan, Kazakhstan, Kyrgyzstan, Turkmenistan and Tajikistan.* London/ Ann Arbor, MI: C. Hurst & Co./The University of Michigan Press.

Lewis E G (1972). *Multilingualism in the Soviet Union*. The Hague: Mouton.

Lewis G (1999). *The Turkish language reform: a catastrophic success*. Oxford: Oxford University Press.

Lo Bianco Joseph (1987). *National policy on languages*. Canberra: Australian Government Publishing Service.

Lo Bianco J & Rhydwen M (2001). 'Is the extinction of Australia's indigenous languages inevitable?' In Fishman J A (ed.). 391–422.

Lo Bianco J & Wickert R (eds.) (2001). *Australian policy activism in language and literacy.* Canberra, Australia: Language Australia.

May Stephen (2001). *Language and minority rights: ethnicity, nationalism, and the politics of language*. Harlow: Essex, England. New York: Longman.

McCarty T L (2002). *A place to be Navajo: Rough Rock and the struggle for self-determination in indigenous schooling*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mesthrie R (ed.) (2002). *Language in South Africa*. Cambridge: Cambridge University Press.

Moys A (ed.) (1998). *Where are we going with languages?* London: Nuffield Foundation.

Myers-Scotton C (1993). *Social motivations for codeswitching: evidence from Africa.* Oxford: Clarendon Press.

Nesiah K (1954). *The mother tongue in education and a language policy for Ceylon*. Colombo: Ola Book.

Neustupný J V (1970). 'Basic types of treatment of language problems.' *Linguistic Communications 1,* 77–98.

Neustupný J V (1989). 'Language purism as a type of language correction.' In Jernudd B & Shapiro M J (eds.) *The politics of language purism*. Berlin/York: Mouton de Gruyter. 211–223.

NicShuibhne N (2001). 'The European Union and minority language rights.' *MOST journal of multicultural studies 3(2).*

Nyati-Ramahobo L (2000). 'The language situation in Botswana.' *Current Issues in Language Planning 1(2),* 243–300.

Omar A H (1998). 'Language planning and image building: the case of Malay in Malaysia.' *International Journal of the Sociology of Language 130,* 49–65.

Paulston C B (1994). *Linguistic minorities in multilingual settings: implications for language policies*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Paulston C B (2004). *Preliminary thoughts on extrinsic linguistic minorities*. (LAUD Paper 605). Essen, Germany: University of Duisburg-Essen.

Pauwels A (1998). *Women changing languages*. New York: Addison Wesley Longman.

Pedersen I L (2003). 'Traditional dialects of Danish and the de-dialectization 1900-2000.' *International Journal of the Sociology of Language 159,* 9–28.

Phillipson R (2003). *English-only Europe?: challenging language policy.* London: Routledge.

Prague School (1973). 'General principles for the cultivation of good language'. Garvin Paul L (trans.) In Rubin J.

& Shuy R (eds.) *Language planning: current issues and research*. Washington, D.C.: Georgetown University Press. 102–111.
Price G (ed.) (2000). *Encyclopedia of the languages of Europe*. London: Blackwell.
Rahman T (2002). *Language, ideology and power: language-learning among the Muslims of North India and Pakistan.* Karachi: Oxford University Press.
Rajagopalan K (2002). 'National languages as flags of allegiance, or the linguistics that failed us: a close look at emergent linguistic chauvinism in Brazil.' *Journal of Language and Politics 1(1),* 115–147.
Roca A (ed.) (2000). *Research on Spanish in the United States: linguistic issues and challenges.* Somerville, MA: Cascadilla Press.
Rubin J & Jernudd B (1979). *References for students of language planning.* Honolulu: East-West Center.
Salhi K (2002). 'Critical imperatives of the French language in the francophone world: colonial legacy – postcolonial policy.' *Current Issues in Language Planning 3(3),* 317–345.
Scharer R & North B (1992). *Toward a common European framework for reporting language competency.* (Position paper). Washington, D.C.: National Foreign Language Center.
Sibayan B (1974). 'Language policy, language engineering and literacy in the Philippines.' In Fishman J A (ed.) *Advances in language planning.* The Hague: Mouton. 221–254.
Skutnabb-Kangas T (2000). *Linguistic genocide in education, or worldwide diversity and human rights?* Mahwah, NJ: L. Erlbaum Associates.
Smolicz J J & Secombe M (1985). 'Community languages, core values, and cultural maintenance: the Australian experience with special reference to Greek, Latvian, and Polish groups.' In Clyne M (ed.) *Australia: meeting place of languages.* Canberra, Australia: Australian National University.
Spolsky B (2002). 'Prospects for the survival of the Navajo language: a reconsideration.' *Anthropology & Education Quarterly 33(2),* 1–24.
Spolsky B (2003). Reassessing Maori regeneration. *Language in Society 32(4), 553–578.*
Spolsky B (2004). *Language policy.* Cambridge: Cambridge University Press.
Spolsky B & Shohamy E (1999). *The languages of Israel: policy, ideology and practice.* Clevedon: Multilingual Matters.
Suleiman Y (2003). *The Arabic language and national identity.* Edinburgh: Edinburgh University Press. Washington, D.C.: Georgetown University Press.
Turell M Teresa (ed.) (2001). *Multilingualism in Spain.* Clevedon, England/Buffalo, NY: Multilingual Matters Ltd.
van Ek J A (1975). *The threshold level.* Strasbourg: Council of Europe.
van Els Theo J M (1992). *The Dutch national action programme on foreign languages.* (English translation): Ministry of Education and Science.
Van Els T J M (2001). 'The European Union, its institutions and its languages: some language political observations.' *Current Issues in Language Planning 2,* 311–369.
Vikor L S (2000). 'Northern Europe: languages as prime markers of ethnic and national identity.' In Barbour S & Carmichael C (eds.) *Language and nationalism in Europe.* Oxford & New York: Oxford University Press. 105–129.
Vikor L S (2001). *The Nordic languages. Their status and interrelations.* Oslo: Novus Press.
Vilela M (2002). Reflections on language policy in African countries with Portuguese as an official language. *Current Issues in Language Planning 3(3),* 306–316.
Weinstein B (1989). 'Francophonie: purism at the international level.' In Jernudd B & Shapiro M J (eds.) *The politics of language purism.* Berlin/New York: Mouton de Gruyter. 53–80.
White D (1997). 'The position and roles of minority languages and their writing systems in China.' *International Journal of the Sociology of Language 97,* 47–57.
Zhou M (2003). *Multilingualism in China: the politics of writing reforms for minority languages 1949–2002.* Berlin: Mouton de Gruyter.
Zhou M (ed.) (2004). *Language policy in the People's Republic of China: theory and practice since 1949.* Dordrecht: Kluwer Academic Publishers.

Language Policies: Policies on Language in Europe

R de Cillia and B Busch, University of Vienna, Vienna, Austria

Theoretical Framework, Terminology

Language Policies, Language Politics, Policies on Language

Questions of language politics and policy on language are always present in society, even when one is not aware of them: it is immaterial whether it is a question of international communication or conflicts to do with minority languages, new minorities and migration, gender-neutral formulations or politically correct use of language. Nevertheless, scientific interest in language policies is relatively young, and the research on language policies has not resulted in the development of any unified theory so far. Even in the area of terminology, there are many ambiguities (Labrie, 1996: 828), complicated by the fact that the

& Shuy R (eds.) *Language planning: current issues and research*. Washington, D.C.: Georgetown University Press. 102–111.
Price G (ed.) (2000). *Encyclopedia of the languages of Europe*. London: Blackwell.
Rahman T (2002). *Language, ideology and power: language-learning among the Muslims of North India and Pakistan*. Karachi: Oxford University Press.
Rajagopalan K (2002). 'National languages as flags of allegiance, or the linguistics that failed us: a close look at emergent linguistic chauvinism in Brazil.' *Journal of Language and Politics 1(1),* 115–147.
Roca A (ed.) (2000). *Research on Spanish in the United States: linguistic issues and challenges*. Somerville, MA: Cascadilla Press.
Rubin J & Jernudd B (1979). *References for students of language planning*. Honolulu: East-West Center.
Salhi K (2002). 'Critical imperatives of the French language in the francophone world: colonial legacy – postcolonial policy.' *Current Issues in Language Planning 3(3),* 317–345.
Scharer R & North B (1992). *Toward a common European framework for reporting language competency.* (Position paper). Washington, D.C.: National Foreign Language Center.
Sibayan B (1974). 'Language policy, language engineering and literacy in the Philippines.' In Fishman J A (ed.) *Advances in language planning*. The Hague: Mouton. 221–254.
Skutnabb-Kangas T (2000). *Linguistic genocide in education, or worldwide diversity and human rights?* Mahwah, NJ: L. Erlbaum Associates.
Smolicz J J & Secombe M (1985). 'Community languages, core values, and cultural maintenance: the Australian experience with special reference to Greek, Latvian, and Polish groups.' In Clyne M (ed.) *Australia: meeting place of languages*. Canberra, Australia: Australian National University.
Spolsky B (2002). 'Prospects for the survival of the Navajo language: a reconsideration.' *Anthropology & Education Quarterly 33(2),* 1–24.
Spolsky B (2003). Reassessing Maori regeneration. *Language in Society 32(4), 553–578.*
Spolsky B (2004). *Language policy.* Cambridge: Cambridge University Press.
Spolsky B & Shohamy E (1999). *The languages of Israel: policy, ideology and practice*. Clevedon: Multilingual Matters.
Suleiman Y (2003). *The Arabic language and national identity.* Edinburgh: Edinburgh University Press. Washington, D.C.: Georgetown University Press.
Turell M Teresa (ed.) (2001). *Multilingualism in Spain.* Clevedon, England/Buffalo, NY: Multilingual Matters Ltd.
van Ek J A (1975). *The threshold level*. Strasbourg: Council of Europe.
van Els Theo J M (1992). *The Dutch national action programme on foreign languages.* (English translation): Ministry of Education and Science.
Van Els T J M (2001). 'The European Union, its institutions and its languages: some language political observations.' *Current Issues in Language Planning 2,* 311–369.
Vikor L S (2000). 'Northern Europe: languages as prime markers of ethnic and national identity.' In Barbour S & Carmichael C (eds.) *Language and nationalism in Europe*. Oxford & New York: Oxford University Press. 105–129.
Vikor L S (2001). *The Nordic languages. Their status and interrelations*. Oslo: Novus Press.
Vilela M (2002). Reflections on language policy in African countries with Portuguese as an official language. *Current Issues in Language Planning 3(3),* 306–316.
Weinstein B (1989). 'Francophonie: purism at the international level.' In Jernudd B & Shapiro M J (eds.) *The politics of language purism*. Berlin/New York: Mouton de Gruyter. 53–80.
White D (1997). 'The position and roles of minority languages and their writing systems in China.' *International Journal of the Sociology of Language 97,* 47–57.
Zhou M (2003). *Multilingualism in China: the politics of writing reforms for minority languages 1949–2002.* Berlin: Mouton de Gruyter.
Zhou M (ed.) (2004). *Language policy in the People's Republic of China: theory and practice since 1949.* Dordrecht: Kluwer Academic Publishers.

Language Policies: Policies on Language in Europe

R de Cillia and B Busch, University of Vienna, Vienna, Austria

Theoretical Framework, Terminology

Language Policies, Language Politics, Policies on Language

Questions of language politics and policy on language are always present in society, even when one is not aware of them: it is immaterial whether it is a question of international communication or conflicts to do with minority languages, new minorities and migration, gender-neutral formulations or politically correct use of language. Nevertheless, scientific interest in language policies is relatively young, and the research on language policies has not resulted in the development of any unified theory so far. Even in the area of terminology, there are many ambiguities (Labrie, 1996: 828), complicated by the fact that the

field, to some extent, has a different terminological structure in different languages.

The beginning of scientific reflection on language policies and politics (LPP) as an independent field was undertaken particularly in the 1960s (Ricento, 2000: 10; Labrie, 1996: 828), in publications such as Haugen (1966), Kloss (1966) or Fishman (1968), and in these the focus was initially on the aspect of language planning. Ricento (2000) distinguishes a total of three stages in the evolution of language policy and planning as an area of research: what is characteristic of early work is the interest in decolonization and state formation, the predominance of structuralism in social sciences and of pragmatism as a strategic orientation. The second phase focused – under the influence of critical sociolinguistics – on social, economic, and political effects of language contact. In the 1990s, linguistic human rights became a strategic aim; postmodern and critical theories foreground language ideologies. On the contemporary situation he comments: "whether the ecology of languages paradigm emerges as the most important conceptual framework for LPP research remains to be seen" (Ricento, 2000: 22).

In any event, the conceptualizations and theoretical approaches of different scholars are different, and the boundaries they draw between their own and related questions are fluid or dependent on their particular conceptualization. Language policy is sometimes seen, for example, as a subdivision of political linguistics (*Politische Sprachwissenschaft*; Maas, 1989). Language and politics are, at any rate, two overlapping subjects. Terms and phrases such as 'social security scroungers,' 'economic migrants,' 'ethnic cleansing' or – to cite examples from the Second Iraq War – 'intelligent weapons,' 'smart bombs,' 'collateral damage' convey a specific approach to reality and are, to some extent, consciously created for this purpose. In a sense, these are also measures of language policies.

A first approximation to a definition of language policy is as follows: language policies deal with political measures targeted at an individual language (e.g., the prohibition to use certain terms) or the relations between different languages, their social importance, function, relevance in international communication, etc. It is therefore a question of an interdisciplinary area of research, and a theory of LP (which has still to be developed) would have to embrace such disciplines as (socio)linguistics, political science, sociology and history, and jurisprudence.

On the one hand, therefore, the object area includes phenomena and measures that are directed at the internal system of a *single* language, politically rule-governed language use, and on the other hand, political phenomena in respect of the status and social functions of languages. It may involve conscious or deliberate actions, or naturally arising developments that are not consciously controlled by politics. And it also includes the level of theoretical concepts and underlying linguistic ideologies with regard to languages, questions of linguistic political theory, and language planning (language planning theory), but also questions of the concrete implementation of such concepts in practice (language planning practice; Fishman, 1994) by means of more or less conscious language-policy measures. Different languages have different conceptual structures for this area: this latter differentiation, therefore, does not exist in German, whereas in English the former is called language politics and the latter language policy or policies; this corresponds to the distinction in French between *la politique linguistique* and *une politique linguistique* (Labrie, 1996: 828–829). Labrie adds a third analytical level of *idéologie linguistique* 'linguistic ideologies,' but this usually remains implicit. The differentiation between measures for the political regulation of language use and phenomena related to the social function of languages is conceptualized in German by most authors as *Sprachpolitik* versus *Sprachenpolitik*.

A number of authors use 'language policy' to refer only to deliberate political measures (e.g., Labrie, 1996), while for other authors it is also important to deal with nondeliberate phenomena, or laissez-faire politics (Ricento, 2000). Finally, language planning (in French *planification linguistique*, but also *aménagement linguistique*, and in German *Sprachplanung*) is sometimes construed as the broader and more comprehensive area (cf., the authors listed by Labrie, 1996: 830), while others treat language policy as the superordinate level. For example, Christ (1991) or Ricento (2000), who says: "I deliberately use 'Language policy' as a superordinate term which subsumes, 'language planning.' Language policy research is concerned not only with official and unofficial acts of governmental and other institutional entities, but also with the historical and cultural events and processes that have influenced, and continue to influence, societal attitude and practices with regard to language use, acquisition and status" (Ricento, 2000: 23).

The most important question to be addressed today by researchers is, according to Ricento: "Why do individuals opt to use (or cease to use) particular languages and varieties for specified functions in different domains, and how do these choices influence – and how are they influenced by – institutional language policy decision-making (local to national and supranational)" (Ricento, 2000: 23).

The German scholar and specialist LP researcher Herbert Christ defines language policy (in the sense of German *Sprachenpolitik*) as any public influence on the communicative radius of languages: "the sum total of all those political initiatives 'from above' and 'from below,' by means of which a particular language or languages are supported in their public currency, their functioning and their dissemination. Like all policies it is subject to conflict, and must be permanently reordered through constant discussion and constant debate" (Christ, 1991: 55). This broad definition of language policy, which we subscribe to, comprises aspects such as language promotion, language maintenance, language conflicts, language wars, language enforcement, linguistic imperialism, and linguistic colonialism.

Language policy may appear in very different forms: as a consciously planned policy in relation to linguistic phenomena, resulting in laws and political measures (e.g., legislation on minorities in European states); as policy on languages that is implicit in existing laws and regulations (e.g., educational laws, curricula, consumer protection laws); as laissez-faire politics that simply let things go and intervene in cases of virulent linguistic conflict; as language policy obscurantism, which simply denies the existence of the phenomena.

Similarly, the area of language policy research in its broadest sense incorporates the phenomena that are labeled *Sprachpolitik* in German, such as politically regulated language use, or the enforcement of measures that serve to exert political control over language use. One extreme example of this was the political creed of National Socialism and its terminological justification in the form of the racist terminology that followed the Nürnberg Race Laws ('Jew,' 'Half-Jew,' German *Jude*, *Halbjude*, etc.).

Language Planning

Language planning is understood to refer to political activities that are used in an attempt to exert conscious and targeted influence on systems of social communication, to introduce desirable changes or to avoid undesirable changes. In other words, it refers to a conscious practical language policy. Here a distinction is made between two areas: corpus planning and status planning. Corpus planning is understood to refer to activities related to the quality and productivity of a language, or its corpus. This includes all measures to do with codification and the setting of norms, which serve to create, or maintain and develop, a standard language, such as questions of the written form, or the development of scientific and literary language. In connection with the development and further expansion of a language, particularly in the area of terminology, the term 'language development' is also used. 'Acquisition planning,' a term used by some scholars, is concerned with policy measures that regulate language teaching and learning.

'Status planning' is understood to refer to those activities that relate to the status of a language: language prestige, the consumer value or market value of a language in international or national communication. Particular languages are preferred in communication or given official backing in their function: for example, at the national level as the state language, official language, and so on (with regulations according to which particular languages are used in offices, schools, law courts, town name boards, and the like); or at an international level as the language of diplomacy and international treaties or as the working language in international organizations such as the United Nations or the Council of Europe.

In the context of status planning the term 'language expansion policy' is also used, and this refers to the spread of a particular language or standard as opposed to other varieties at the national level, or to the spread of a particular language as opposed to other languages at the international level. An example of the former would be the promotion in France of the dialect of the *Île de France*, the *langue d'oïl,* as the French national language, at the expense of other varieties, from the 16th century onwards. This happened in particular at the expense of the *langue d'oc*, which had been the language of a highly developed culture in the Middle Ages. Language dissemination policy, or Language promotion policy, at an international level aims primarily at promoting the status of one language with regard to other nations or the official languages of other nations. One central agency in language expansion at the national level is the school, which plays a decisive role in the spread of standard languages and national languages, but also in the suppression of minority languages and the spread of particular foreign languages rather than others. At the international level, the cultural institutes play an important role in the implementation of the external language policies or expansion policies of particular nations (e.g., United States Information Agencies/Services, British Council, Institut Français, Goethe Institut, Österreich Institut, Instituto Cervantes, Istituto Italiano) (cf., Christ and de Cillia, 2003).

The Analysis of National and International Language Policies

Traditionally, a distinction is made between national and international language policies, although of

course the boundaries between are also fluid. To analyze LP on a national level, one should investigate the following questions: What language policy and language planning measures, and what language laws are set up in relation to the role, meaning, and status of languages that are spoken by the inhabitants of a particular state (as first, second, or foreign languages)? This includes all measures that affect the standardization, dissemination, and promotion of the relevant languages, within but also beyond the territory of the state in question. The first task, therefore, would be to investigate which languages, varieties, and codes are used by the population living in the territory of a particular state. What status do the languages in question have in the particular society? What forms of bilingualism and diglossia exist, and what is the functional distribution of the individual languages in diglossic and/or multilingual situations? What language conflicts exist?

The analysis of language attitudes needs special attention, as attitudes toward languages/language varieties, language use and toward the speakers of particular languages/varieties allow extrapolations on expectations, acceptance, and efficiency concerning language policies (Garret *et al.*, 2003; *see also* **Language Attitudes**).

Another important question is the relationship between language policy and language legislation: what linguistic regulations are in existence (in the constitution, consumer protection laws, educational laws, laws controlling the media, regulations on official and legal language(s), military laws, topographical notices, etc.)? Further, there must be an investigation into whether there is a coordinated national language policy that provides for a development and implementation of the legal requirements on language according to scholarly standards, and whether there is any body that is responsible for the development of an explicit language policy (political body, panel of experts, or a similar organization). One must also ask what contradictions there are between linguistic rights guaranteed by law and language policy in practice. In principle, all political fields may be analyzed with this in mind: i.e., education and school policy (medium of instruction, minority languages and foreign languages in the curriculum), media policy, consumer policy, cultural policy, foreign policy, migration policy, and so on. Finally, it is necessary to analyze the measures implemented in a particular state on language planning: what corpus planning and status planning measures are implemented/not implemented in respect of the languages in question. Are there academies as bodies that set norms, are there dictionaries, grammars, codices for the particular languages, are there measures for terminological development (e.g., in the areas of modern technology, the Internet)? With regard to status planning, one should ask what measures are taken there to enhance or guarantee the consumer value of the particular languages on the national and international language markets?

Concerning the analysis of international language policies, a first step is the analysis of measures that have been incorporated into the foreign and cultural policy of individual states to safeguard the role of languages in international and supranational communication, i.e., as languages of regional, supraregional, and international communication, as foreign languages in other countries, as working languages in international organizations, at bilateral or multilateral international meetings, as the language of international treaties, and so on.

Beyond the analysis of the above-mentioned questions of international policies, the international status of languages is of the utmost importance. It can be characterized by means of a number of different factors. The first, of course, is the number of speakers who have a particular language as their first language. (Here Mandarin Chinese comes before Hindi, English, Spanish, Arabic, Bengali, Portuguese, Russian, Japanese, German, and French, according to Grimes (2004) or as a second language (Chinese before English, Spanish, Hindi, Russian, French, German, Portuguese, again according to Grimes (2004). No figures are given for Arabic). Other criteria are the importance of languages in politics and diplomacy (with English as national or official language in 58 countries, French 32, Spanish 22, German eight, Portuguese seven (Grimes, 2004). No figures are given for Arabic, but it presumably has this status in some 20 states.) As a working language, we find English in 16 international organizations, French in 12, Spanish in nine, Russian in five, Arabic in two and German in one organization (Ammon, 1999: 111). Further criteria are the economic importance of languages that can be expressed, for example, in their economic strength, in the Gross Social Product that members of a particular language community achieve worldwide. Here, English is ahead of Japanese, German, Spanish, and French (Ammon, 1999: 101); or their role as academic languages (in 1996, 90.7% of all publications in the natural sciences were in English, 2.1% in Russian, 1.7% in Japanese, 1.3% in French, and 1.2% in German (Ammon, 1999: 109); and as an Internet language. One source estimates the number of English web pages at 56.4%, German 7.7%, French 5.6%, Japanese 4.9%, Spanish 3.0%, and Chinese 2.4%. The statistics for languages used for searches (using the Google search engine) are as follows:

English 57%, German 12%, Japanese 7%, Spanish 6%, French 5%, and Chinese 3% (StADaF, 2003: 4).

Different sources give somewhat different figures for all these criteria, but the relationships are always more or less the same. About a dozen major languages, with 100 million or more speakers, play an important role in international language politics, English plays a dominant role as a global language in all areas, and in politics and diplomacy French, which in relation to the number of its speakers, still has great significance.

A central role in the maintenance, improvement, or protection of the status of a language at the international level belongs to the language expansion policy outlined above. Several states therefore maintain cultural institutes in many parts of the world. For instance, there are 200 US Information Agencies in 143 different countries, 160 British Council centers in 109 countries, 151 *Instituts Français* in 92 countries, 126 *Goethe Institute* in 76 countries, 90 *Istituti Italiani di Cultura* in 61 countries and 40 *Institutes Cervantes* in 20 countries (Müller, 2004: 152). France – the only nation in the world – even has an independent department within the foreign ministry that is responsible for external language policy (*Direction Générale de la Cooperation Internationale et du Développement*).

The boundaries between language dissemination policy and linguistic imperialism (Phillipson, 1992) are fluid, and this policy is justified by various ideologies such as the supposedly natural superiority of a particular language, the ideology of a neutral lingua franca (claimed, for example, for English) or the special clarity of a language (*clarté de la language française*). For French, a particular feature of all French-speaking countries (*la francophonie*) is construed discursively (and also politically and economically), and more recently the concept of plurilingualism is being promoted at an international level, since English can no longer be denied the role of **the** international lingua franca.

Trends and Tendencies in Inter/national Language Policies

Present-day developments in language policy are contradictory: on the one hand there is hegemonization and on the other hand differentiation. The figures given above show clearly that English has taken on an ever more dominant role in connection with processes of globalization and processes of political, economic, and cultural hegemonization emanating from the United States.

In the context of globalization, the competences of the nation–state are shrinking at an increasing rate. It is surrendering power to international or supranational bodies, and at the same time allowing new media technologies (Internet, satellite TV, etc.) to create new communication spaces that are less linked to territory than before. Moreover, regional supranational groups and bodies such as the European Union have also arisen together with globalization. Although language policy in the European Union is an area determined by the subsidiarity principle, which leaves relative autonomy to the individual member states, different common regulations and programs exert a considerable influence. On the other hand, cities and regions on a substate level also have gained autonomy in the shaping of their particular social and cultural policies. The process of de-centering the nation–state as a central principle of political organization also makes it necessary to rethink questions of language policy. Here, particular attention has to be given to such market-conditioned language policy players as the language learning, culture, and media industries that previously occupied a less central position (Labrie, 2003; Wright, 2004; Busch, 2004).

In a kind of counter-movement, there are attempts on the part of some states to affirm the role of the nation–state in the area of language policy, for example through strengthening the constitutional role of national languages or through the increasing requirement in European countries for immigrants to be able to demonstrate knowledge of the state language before being granted citizenship, or even the right of permanent residence.

The central role that language policy still plays in the process of affirmation of nation states or in the construction of national identities is shown in the example of former Yugoslavia, where the break-up of the federal state into individual states was accompanied by the deconstruction of the umbrella language of Serbo-Croatian and the establishment of the three languages Serbian (Serbo-Croatian), Croatian, and Bosnian as state languages in the three respective states (Bugarski, 2004). There were also processes of differentiation in language policy in the Republic of South Africa where, after the end of the racist apartheid regime, there was an upgrading of African languages and sign language. The establishment of nine African languages as official languages and a comprehensive commitment to a policy of multilingualism are viewed as a central element in the process of social transformation (Alexander, 2001). Further questions of international language policy will be discussed below, using the example of language policy in Europe.

Language Policies in Europe

Languages in Europe

Of the estimated 2500–8000 languages in the world, only a few hundred (230; Grimes, 2004) are spoken in Europe, and laws concerning languages normally distinguish between national languages, autochthonous minority languages such as Basque or Gaelic, new languages arising from immigration, such as Turkish or Arabic, and sign languages.

Among the national languages, after English and French the most important European languages are Russian (158 million speakers) and German (121 million) (Crystal, 1997: 289). Before May 1, 2004, 24% of EU citizens spoke German, 16% English, French or Italian, and 10% Spanish (Eurobarometer, 2000: 93). Of German speakers, incidentally, 96% live in Europe, compared with 47.6% of French speakers, 10.7% for English, 11.2% for Spanish, and 5.4% for Portuguese (StADaF, 2003: 4). If one includes foreign language knowledge, the following picture emerges: English is spoken by 47%, German 32%, French 26%, Italian 18%, and Spanish 14% (Eurobarometer, 2000: 94).

On the other hand, it has been estimated that in the great European metropolises, approximately one-third of the population conduct their daily lives not only in the national language, but also in a range of other languages. Twelve percent of the 370 million EU citizens use 40–60 (depending on various sources) different minority languages (Krausneker, 2003: 34).

Official Languages and Language Policies in Europe

For international and supranational communication, there are in principle two different models: the lingua franca/dominant language model, according to which a lingua franca (e.g., English, French, Esperanto) serves as a means of communication between different language communities, and a model of linguistic pluralism, or linguistic diversification, according to which as many different languages as possible are used as means of communication. In Europe, a pluralistic model of communication is pursued, at least among theoreticians, and European multilingualism is accepted as an important element in the European identity.

The basic principle for the institutional language policy in the EU was determined by the Ordinance no. 1 of April 15, 1958 (Language Charter of the EU). It stipulates that the various national languages of the member states are simultaneously "the official and working languages of the organs of the community." (There are currently 20 languages: Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Slovak, Slovenian, Spanish, and Swedish.) The ordinance refers to written documents, and also requires that "ordinances and other written documents of general validity" be published in all official languages, and that the bulletin of the community appear in all of the official languages. To implement this policy, the EU maintains language services in all the different organs (Parliament, Council, Commission, and so on) in which – before the most recent expansion on May 1, 2004 – a total of some 4000 translators and interpreters were employed (language staff make up around 12.6% of total staff). Apart from the interpreting and translation service, as further measures for the promotion of multilingualism the EU has developed special programs at three different levels: assistance in machine translation (e.g., Systran), terminology databases (e.g., EURODICAUTOM, EUTERPE), and promotional programs for language learning such as LINGUA, SOKRATES, and LEONARDO DA VINCI.

The analysis of actual communication in the EU demonstrates that the policy of multilingualism is most fully implemented at the level of the internal working languages in the European Parliament, and then in the European Court of Justice in external communication. (The internal language here is French.) In the European Council and Commission, most work and negotiation is done in English and French, and to a lesser extent in German. Here English is becoming ever more important as the internal working language at the expense of French: in 1986, for example, 58% of the Commission's primary texts were written in French, 26% in English, 11% in German and 5% in other languages. In 1997, the figures were as follows: English 45.4%, French 40.4%, German 5%, other languages 9.2% (Truchot, 2001: 27).

The translation service and its costs are often advanced as an argument in favor of a dominant language communication model, where English would be adopted as the only internal working language of the EU. On the other hand, language policy experts consider the (political) costs of monolingualism to be too high and plead for a more conscious and clearly defined EU policy of multilingualism (cf., Phillipson, 2003; de Cillia *et al.*, 2003, and the contributions documented therein.)

Language Policies in Education

A central role in European language policy is played by school language policy, since foreign language

competence is acquired first and foremost in the school system (59% of Europeans gained their knowledge of a first foreign language in their secondary school (Eurobaromètre, 2001: 23)). In the year 2000, approximately half the population throughout the EU had no foreign language knowledge at all (an average of 47% throughout the EU, while 26% of Europeans spoke two foreign languages). The most important foreign language is English (41%), ahead of French (19%), German (10%), Spanish (7%), and Italian (3%) (Eurobaromètre, 2001: 2). This has a direct relationship to the foreign languages that are taught in schools: in 1999–2000, 87% of students in secondary schools in 27 European countries were learning English, 25% French, 15% German, and 7% Spanish (European Commission *et al.*, 2002: H-17). In the new EU member states, the language teaching policy is probably diversifying more strongly than in the old member states. For example in the Czech Republic, in the 1998–1999 academic year, 390 518 pupils were learning English, 344 247 German, and 8744 French in elementary schools (age range, 6–15) (Nekvapil, 2003: 78).

Research into language policy and language learning in the last few decades has produced a series of proposals for intensifying and diversifying foreign language teaching in Europe, and have made it their objective that pupils should learn two foreign languages during their compulsory education and three in postcompulsory education. To implement a foreign language policy leading to multilingualism, specialists have proposed a whole range of measures at the levels of school organization and teaching methodology. For instance, foreign language teaching should, in general, start early; use of the foreign language as a medium of instruction should increase; curricula should be more flexible; passive multilingualism should be encouraged, and multilingualism available in classrooms should be exploited, involving the languages of both old and new minorities (see Besters-Dilger *et al.*, 2003: 282; de Cillia *et al.*, 2003: 11).

Minority Languages

The majority of Europe's legal systems recognize within the borders of particular states a number of minority languages. Which languages are recognized as minority languages, how their status is defined, and in which domains these languages may be used varies considerably from country to country and depends very much on particular political constellations. There is also a considerable fuzziness concerning terminology in the field: in some contexts, minority languages are referred to as regional languages, in others as lesser used languages or as autochthonous minority languages.

The Council of Europe's Charter for Regional or Minority Languages (Council of Europe, 1992), which came into force in 1998, is the first legally binding and individually enforceable international instrument in the field of language policy. On the subject of minority languages, the Charter marks a change of paradigms concerning the attitude toward minority languages: whereas earlier international treaties protected speakers of minority languages only through nondiscrimination clauses limiting state influence, the Charter puts the ratifying states under an obligation to give active protection and support for minority languages. In 2004, only 17 of the 45 member states of the Council of Europe had ratified the Charter. The European Union has established an Office for Regional and Lesser Used Languages, which protects the interests of speakers of minority languages within the EU.

In the 1970s, minority languages became an important topic in sociolinguistic research when, in a counter-move to the previous centralizing tendencies, regionalization was not only supported at the political level, but also by social movements that began to demand rights for minorities.

A new approach to borders in Europe, resulting from the end of the bi-polar division in the world, the process of European integration, and globalization in the world economy, have, since the 1990s, not only brought about new minority–majority configurations, but have also triggered a rethinking of language policies in the field. Since that time, greater interest has been taken in the rights of speakers and the development of nonterritorial languages, particularly Romany. Other minority languages are increasingly seen as regional languages that can function as lingua francas in cross-border situations. Frequently, such functions are not fulfilled by national standard languages, which tend to emphasize differences, but by regional dialects or vernaculars that are intelligible on both sides of a border and that are therefore undergoing a process of revival (Raasch, 2002: 205).

The interest in migrant languages began in the domain of language in education when it became obvious in Western Europe that migration could not be considered a temporary phenomenon. Early research within the multicultural paradigm was mainly concerned with language acquisition and bi- or multilingual education. For the most part, migration and mobility are no longer seen as temporary phenomena but as a consequence of the process of globalization. Research interest has shifted in recent years and

focuses more on (hybrid) youth codes, and on languages in cultural expression and in the media. Although it is claimed that linguistic diversity should be considered a resource, migrant languages remain for the moment low-status languages in most cases, and there are no linguistic rights granted to their speakers, although some of these linguistic groups are statistically more important than autochthonous minorities.

One group that is even more victimized within the EU than speakers of minority languages is users of sign languages (SL): no official document deals with the right of SL users, and no official statistics give any information about them. Deaf people are seen as disabled, whereas they consider themselves to be a linguistic/cultural minority. For example, the EUD (European Union of the Deaf), their central representative organization in Europe, regards the aspect of being handicapped rather as a lack of access to resources. The number of deaf sign language users has been estimated at nearly half a million people (Krausneker, 2003: 15). Resolutions in the European Parliament (from 1988 and 1998) for the recognition of sign languages have so far had little effect (at present sign languages are recognized in the following European countries: Denmark, Finland, France, Germany, Greece, Norway, Portugal, and Sweden). In educational policy, oral models still dominate, rather than a bilingual form of instruction that would support language acquisition.

Recent documents by the European Union and the Council of Europe have begun to question the differentiation into autochthonous and allochthonous minority languages. In this sense, the action plan promoting language learning and linguistic diversity, an EU 2004–2006 action plan (Commission of the European Communities, 2003) is now explicitly opening programs that encompass a language policy dimension (such as the above-mentioned SOCRATES Program), not only for the official languages of member states and lesser used EU languages, but also for the languages of migrants, sign languages, and the languages of the EU's most important trade partners.

See also: Arabic; Basque; Bilingual Education; Bilingualism; Discrimination and Language; Education in a Former Colonial Language; Education in a Multilingual Society; Educational Linguistics; English: World Englishes; Foreign Language Teaching Policy; French; German; Identity and Language; Immigrant Language Education; Italian; Language Attitudes; Language Education for Endangered Languages; Language Education of the Deaf; Language Education Policies in North America; Language Education Policy in Europe; Language Ideology; Language Loyalty; Language of Tertiary Education; Language Planning and Policy: Models; Language Politics; Language Spread; Law and Language: Overview; Law on Language; Lingua Francas as Second Languages; Linguistic Decolonization; Linguistic Rights; Media and Language: Overview; Migration and Language Planning; Migration and Language; Minorities and Language; Minority Languages: Oppression; Mother Tongue Education: Nonstandard Language; Mother Tongue Education: Standard Language; Multiculturalism and Language; Multilingualism: Pragmatic Aspects; Politics and Language: Overview; Politics of Teaching; Pragmatics: Linguistic Imperialism; Russian; Second and Foreign Language Learning and Teaching; Serbian–Croatian–Bosnian Linguistic Complex; Sign Language: Acquisition; Sign Language: Communities and Cultures; Sign Language: History of Research; Society and Language: Overview; Spanish; Standardization; Teaching of Minority Languages; Turkish.

Bibliography

Alexander N (2001). *Südafrika. Der Weg von der Apartheid zur Demokratie.* Munich: C. H. Beck.

Ammon U (1999). 'Grundzüge der internationalen Stellung der deutschen Sprache – mit Hinweisen auf neueste Entwicklungen.' *Zeitschrift für Anglistik und Amerikanistik 47(2),* 99–119.

Besters-Dilger J, de Cillia R, Krumm H-J & Rindler-Schjerve R (eds.) (2003). *The future of European multilingualism in the enlarged European Union. Die Zukunft der europäischen Mehrsprachigkeit in einer erweiterten Europäischen Union. L'avenir du multilinguisme européen dans l'Union Européenne élargie.* In Zusammenarbeit mit Gerald Rosskogler. Klagenfurt: Drava.

Bugarski R (2004). 'Language and boundaries in the Yugoslav context.' In Busch B & Kelly-Holmes H (eds.) *Language, discourse and borders. Current issues in language and society.* Clevedon: Multilingual Matters. 21–38.

Busch B (2004). *Sprachen im Disput. Medien und Öffentlichkeit in multilingualen Gesellschaften.* Klagenfurt: Drava.

de Cillia R, Krumm H-J & Wodak R (eds.) (2003). *The cost of multilingualism – globalisation and linguistic diversity. Die Kosten der Mehrsprachigkeit – Globalisierung und sprachliche Vielfalt.* Vienna: Verlag der Österreichischen Akademie der Wissenschaften.

Christ H (1991). *Fremdsprachenunterricht für das Jahr 2000. Sprachenpolitische Betrachtungen zum Lehren und Lernen fremder Sprachen.* Tübingen: Narr.

Christ H & de Cillia R (2003). 'Ausländische Kulturinstitute in den deutschsprachigen Ländern.' In Bausch K, Christ H & Krumm H-J (eds.) *Handbuch Fremdsprachenunterricht,* 4. vollständig neu bearbeitete Auflage. Tübingen: Francke. 601–604.

Commission of the European Communities (2003). 'Promoting language learning and linguistic diversity: an action plan 2004–2006.' Brussels, 24.3.2003.

Council of Europe (1992). *European charter for regional or minority languages.* European Treaty Series Nr. 148. Strasbourg: Council of Europe.

Crystal D (1997). *The Cambridge encyclopedia of language* (2nd ed.). Cambridge: Cambridge University Press.
Eurobarometer (2000). Report Number 52. http://europa.eu.int/comm/public_opinion/standard_en.htm,15.9.2004.
Eurobaromètre (2001). Eurobaromètre 54 Spécial *Les européens et leurs langues*. http://europa.eu.int/comm/public_opinion/standard_en.htm, 15.9.2004.
European Commission/Eurydice/Eurostat (2002). 'Key data on education in Europe 2002. Chapter H: foreign languages.' Brussels.
Fishman J A (1968). 'Sociolinguistics and the language problems of the developing countries.' In Fishman J A, Ferguson C A & das Gupta J (eds.) *Language problems of developing nations*. New York: John Wiley & Sons. 3–16.
Fishman J A (1994). 'Critiques of language planning: a minority languages perspective.' *Journal of Multilingual and Multicultural Development 15*, 91–99.
Garrett P, Coupland N & Williams A (2003). *Investigating language attitudes. Social meanings of dialect, ethnicity and performance*. Cardiff: University of Wales Press.
Grimes B (2004). *Ethologue. Languages of the world*, Web version. http://www.ethnologue.com/, 15.9.04.
Haugen E (1966). *Language conflict and language planning: the case of modern Norwegian*. Cambridge: Harvard University Press.
Kloss H (1966). 'Types of multilingual communities: A discussion of ten variables.' *Sociological Inquiry 36*, 7–17.
Krausneker V (2003). 'Has something changed? Sign languages in Europe: the case of minorised minority languages.' *Deaf Worlds. International Journal of Deaf Studies 19(2)*, 33–46.
Labrie N (1996). 'Politique linguistique.' In Goebl H, Nelde P, Stary Z & Wölk W (eds.) *Kontaktlinguistik. Ein internationales Handbuch zeitgenössischer Forschung, 1. Halbband*. Berlin: de Gruyter. 826–833.
Labrie N (2003). 'Le rôle des langues des pays candidats à l'Union Européenne.' In Besters-Dilger J, de Cillia R, Krumm H-J & Rindler-Schjerve J (eds.) *Mehrsprachigkeit in der erweiterten Europäischen Union*. Klagenfurt: Drava. 202–212.
Maas U (1989). *Sprachpolitik und politische Sprachwissenschaft*. Frankfurt: Suhrkamp.
Müller B (2004). Eine Weltsprache zwischen Tradition und Innovation: Aktuelle Tendenzen französischer Sprachenpolitik in der Welt unter besonderer Berücksichtigung der Umsetzung in Österreich. Ph.D. diss., University of Vienna.
Nekvapil J (2003). 'On the role of the languages of adjacent states and the languages of ethnic minorities in multilingual Europe: the case of the Czech Republic.' In Besters-Dilger J, de Cillia R, Krumm H-J & Rindler-Schjerve R (eds.) *The future of European multilingualism in the enlarged European Union. Die Zukunft der europäischen Mehrsprachigkeit in einer erweiterten Europäischen Union. L'avenir du multilinguisme européen dans l'Union Européenne élargie*. Zusammenarbeit mit Gerald Rosskogler, Klagenfurt: Drava. 76–94.
Phillipson R (1992). *Linguistic imperialism*. Oxford: Oxford University Press.
Phillipson R (2003). *English-only Europe?: Challenging language policy*. London: Routledge.
Raasch A (2002). 'Grenzregionen und die Mehrsprachigkeit.' In Kelz H (ed.) *Die sprachliche Zukunft Europas. Mehrsprachigkeit und Sprachenpolitik*. Baden-Baden: Nomos. 193–209.
Ricento T (2000). 'Historical and theoretical perspectives in language policy and planning.' In Ricento T (ed.) *Ideology, politics and language policies. Focus on English*. Amsterdam: Benjamins. 9–24.
StADaF = Ständige Arbeitsgruppe Deutsch als Fremdsprache (2003). Deutsch als Fremdsprache. Erhebung 2000. Berlin u.a. StADaF. http://www.goethe.de/uun/pub/de18483.htm, 15.9.04.
Truchot C (2001). 'Le français langue véhiculaire en Europe.' In Ammon U, Mattheier K & Nelde P (eds.) *Verkehrssprachen in Europa – außer Englisch/Lingua francas in Europe – except English/Langue véhiculaires en Europa – sans l'anglais*. (Sociolinguistica Nr. 15). Tübingen: Niemeyer. 19–31.
Wright S (2004). *Language policy and language planning. From nationalism to globalisation*. Basingstoke: Palgrave Macmillan.

Relevant Websites

http://www.diplomatie.gouv.fr – Direction Générale de la Cooperation Internationale et du Développement.
http://conventions.coe.int – Council of Europe's Charter for Regional or Minority Languages.

Language Policy in Multilingual Educational Contexts

S Romaine, University of Oxford, Oxford, UK

Global Distribution of Multilingualism

Bilingualism and multilingualism are a normal and unremarkable necessity of everyday life for the majority of the world's population. Linguists estimate that there are roughly 6800 languages in the world, but only about 200 nation-states. With more than 30 times as many languages as there are countries, bilingualism or multilingualism is present in practically every country in the world, whether it is officially recognized or not (Romaine, 1995). This means that in a broad sense multilingual educational contexts can be understood to encompass the educational practices of most countries in the world. The varied cultural and linguistic contexts existing in contemporary societies around the globe pose complex challenges for policy makers in many areas. The centrality of language to education means that policies concerning choice of which language(s) to use as the medium of instruction are essential, even if the need is not always overtly acknowledged. In addition, the need for teaching of additional languages as subjects is widely recognized as schools have a critical role to play in providing the bi- and multilingual skills that have become increasingly necessary in the modern world. In this article the terms 'bilingualism' and 'multilingualism' will be used interchangeably to refer to the routine use of two or more languages in a community (*see* **Bilingualism**).

Despite the near-universal presence of more than one language in every country, the global distribution of linguistic diversity is strikingly uneven. Papua New Guinea alone contains 13.2% of the world's languages, but only 0.1% of the world's population and 0.4% of the world's land area. The overall ratio of languages to people is only about 1 to 5000. If this ratio were repeated in the United States, there would be 50 000 languages spoken there (Nettle and Romaine, 2000). Over 70% of all the world's languages are found in just 20 nation-states, among them some of the poorest countries in the world. They include Papua New Guinea (823), Indonesia (726), Nigeria (505), India (387), Mexico (295), Cameroon (279), Australia (235), the Democratic Republic of Congo (218), Brazil (192), United States (175), the Philippines (169), Malaysia (139), Tanzania (135), Vanuatu (109), Russia (100), Vietnam (93), Laos (82), Ivory Coast (77), Ghana (79), and Solomon Islands (69). These data come from the *Ethnologue*, a database compiled by SIL International.

Need for Language Policy and Planning

The pervasive presence of some degree of multilingualism indicates a universal need for language policy and planning in order to ensure that members of different language groups within a country or other administrative unit have access to and can participate in important societal institutions such as schools, government, and media (*see* **Language Planning and Policy: Models**). Schooling is one of the most critical sites for planning because education is the primary societal institution through which legitimation for the state's dominant language is sought. Formal education is often the first point of contact children have with the world outside their own community. Speakers of languages other than the official and national languages recognized for instructional purposes are often at a disadvantage. The poor school achievement of minority group children due to discontinuities between home and school language is well documented (Corson, 1990; Tollefson, 1995; *see* **Educational Failure**).

The first sociolinguists to tackle questions of language policy and planning were concerned with the language problems of developing nations such as Malaysia and India (Fishman *et al.*, 1968). Many countries under former colonial rule designated the colonial language as their sole official language for education and government. Ivory Coast, for example, declared French as its official language. Some have in addition specified national languages which may be compulsory in education. Others such as Indonesia replaced the former colonial language with their own (*see* **Education in a Former Colonial Language**).

From the 1970s onward, scholars have been concerned with education in migration contexts, particularly in Europe, which has seen the rise of increasingly diverse populations in countries such as Portugal and Iceland, often cited as examples of monolingual nations. The illusion of linguistic homogeneity is belied by the existence of a number of minorities in both these places. Portugal has a population of about 10 000 speakers of Mirandese (Miranda do Douro) concentrated in small villages in the northeastern part of the country. Because both Portuguese and Mirandese are closely related Romance languages, Mirandese has been thought of as a dialect of Portuguese. However, in 1999 the Portuguese parliament recognized it as a regional minority language and has undertaken some steps to protect it. It is optionally taught in some local schools, and work on grammars and dictionaries is under way. Portugal also has large numbers of immigrants from its former colonies (Angola, Mozambique, Guinea-Bissau, Cape Verde,

São Tomé and Principe, East Timor, and Brazil). There may be as many as a quarter of a million 'new' immigrants (many of them illegal) from eastern Europe, especially Ukraine and Russia. Immigrants now comprise about 5% of the population, one of the highest proportions in the European Union, up from less than 2% at the turn of the 21st century. Iceland too has witnessed an influx of immigrants from Asia, especially since the 1990s, in addition to those coming from European countries. Although only around 3% of the population is non-Icelandic in origin, as many as 40 different languages may be spoken in addition to Icelandic.

An even more recent area of concern has been the notion of language rights, as individual and collective rights of persons belonging to linguistic minorities have been increasingly acknowledged in international human rights law and encoded in various legal instruments (May, 2001; Skutnabb-Kangas and Phillipson, 1994).

Components of Language Policy

Spolsky (2004: 5) distinguished three components of language policy: language practices, language beliefs or ideology, and language planning. The notion of language practices concerns the choices members of a community make among the varieties available for use. Consider, for instance, the many Haitians or Cubans who have immigrated to cities such as Miami, who may use English to varying degrees in addition to Haitian Creole French or Spanish. Or consider Saami speakers in Norway, who may know Saami (an indigenous language of northern Norway, Sweden, and Finland), Norwegian, and English to varying degrees. Language beliefs include attitudes toward and beliefs about these varieties. Until recently attitudes toward Saami have been quite negative, among both Saami and non-Saami. Majority populations often show little enthusiasm for the languages of immigrant minorities either, even when the language concerned is a world language such as Spanish (as is the case in the United States) or Arabic (the language of many immigrants in France and the Netherlands). This is due to status differences between the majority and minority populations. Distinctive food, dress, song, etc. are often accepted and allowed to be part of the mainstream, but language much less so. The idea that linguistic rights need protection has never been part of American culture, and so they have not been seen as central to U.S. courts unless allied with more fundamental rights such as educational equity, etc. (Schiffman, 1996).

Language planning includes any efforts to modify practices or beliefs by means of some form of management or intervention. It usually takes the form of a set of planned and managed interventions supported and enforced by law and implemented by official government agencies. Many countries encode language policies of one sort or another in their constitutions, laws, or other official documents. UNESCO sponsors the MOST (Management of Social Transformations) Clearing House on Linguistic Rights, designed to provide information for legislators, decision makers, researchers, and other representatives of both governmental and nongovernmental organizations. The database provides an overview of the most important international legal instruments, major nongovernmental documents, and national constitutions containing provisions relating to language and the rights of linguistic minorities, and a bibliography on linguistic rights in international human rights law (*see* **Law and Language: Overview**). In conjunction with other sources of data (Skutnabb-Kangas, 2000: 297–311), this information can be used for further sociolinguistic analysis as well as for the development of multilingual policies.

The MOST database lists 163 constitutions containing some mention of language; 22 countries either have no constitution at all or have a constitution that contains no provisions relating to language. Perhaps the most common kind of provision is to declare a language or languages as official or co-official, or as a national language. Nevertheless, fewer than 4% of the world's languages have any kind of official status in the countries where they are spoken. The fact that most languages are unwritten, not recognized officially, restricted to local community and home functions, and spoken by very small groups of people reflects the balance of power in the global linguistic marketplace. Around 100 constitutions specify one or more official or national languages with special privileges of use. Seventy-eight mention a single official or national language. The constitution of France says that "the language of the Republic shall be French."

More than 20 countries have more than one official language. India, for instance, has 19 and South Africa has 11. The constitution of Vanuatu states that the national language is Bislama, and the official languages are Bislama, English, and French. The principal languages of education, however, are English and French. India's constitution codifies a variety of provisions protecting linguistic minorities, including the right to establish and administer educational institutions of their choice, and freedom from discrimination on grounds of language. In addition to specifying Hindi as the official language, it grants rights to regional state languages and specifies which languages can be used for communication between states, and between states and the national government. Every

state is supposed to endeavor to provide adequate facilities for instruction in the mother tongue at primary level to children belonging to linguistic minority groups, and there is a provision establishing a Special Officer for Linguistic Minorities. When a language is spoken by 30% or more of the population in any state or district, it is recognized as bilingual and the relevant minority language is placed on the same footing as the regional language for use by public authorities.

In practice, no country gives official status to every single language spoken within its territory. Where language policies exist, they inevitably privilege a limited set of languages. Even where explicit policies do not exist, governments have to operate in some language(s). This means that policy is implicit even if no specific mention is made of language. Here is where an examination of practice is essential. The presence of many languages other than English in industrialized countries such as the United States and Australia often comes as a surprise because these countries have generally operated and seen themselves as largely monolingual English nations, despite the presence of a considerable number of indigenous and (im)migrant communities using other languages. The United States is a classic example of a country with no official language policy. The term 'benign neglect' is sometimes used to describe cases where a state has no codified policy specifying which languages are official. Nevertheless, when a multilingual country uses one or more languages exclusively in public schools, and in the administration of state services and activities, it is making a distinction based on language. In showing a preference for some language(s), whether designated as official or national or not, the state's decision benefits those for whom the chosen language(s) is a primary language, to the detriment or disadvantage of others who either have no or lower proficiency and are denied the benefit or privilege of using their own primary language. The only cases where immigrant and indigenous minorities receive equal treatment are in those countries where neither group is given any special status (Kymlicka and Patten, 2003).

The term '*de facto*' ('by fact') is used for policies that operate covertly, implicitly, without necessarily having any official written support in legal documents. '*De jure*' ('by law') policies are overt, explicit, officially and legally defined. Probably most majority languages dominate in many domains where they have only *de facto* and no legal status. English is the dominant *de facto* or official language in over 70 countries. French has official or co-official status in 29 countries. The majority of countries in the world operate either *de facto* or *de jure* as monolingual states in recognizing only one language for use in education. This does not always mean that no other languages are used in education, but rather that they do not have official status. Again one must look to practice in individual cases to assess the situation.

Language Policies in Nation-States

The nation-state is the most critical unit of analysis because it is the policies pursued within national boundaries that gives some languages (and their speakers) the status of majority and others that of minority language. The term 'minority' is ambiguous because it may have both numerical and social/political dimensions. It is generally a euphemism for nonelite or subordinate groups, whether they constitute a numerical majority or minority in relation to some other group that is politically and socially dominant. What is common to most minority languages from a sociopolitical perspective is the fact that their status is defined in relation to some administrative unit, which in the modern world is generally the nation-state. Mandarin Chinese, with 900 million speakers, is spoken by more people than any other language in the world. In China, it has the status of majority language, but in many other countries such as Malaysia, it is a minority language. Catalan (Catalan-Valencian-Balear) is spoken by a minority of people within Spain, but by a majority in Catalonia, where it has official recognition. A minority language in a large country may be a majority language in a smaller country. Some languages, such as the signed languages used among deaf speakers, are minority languages in all contexts (*see* **Minorities and Language**).

The linguistic heterogeneity of many countries reflects the linguistic arbitrariness of shifting political boundaries that have encapsulated distinct ethnic groups or nationalities with their own languages. All nation-states, whatever their political ideology, have persecuted minorities in the past and many continue to do so today. Many indigenous people today such as the Welsh and the Basques find themselves living in nations that they had no say in creating and are controlled by groups who do not represent their interests and in some cases actively seek to exterminate them. More than 80% of the conflicts in the world today are between nation-states and minority peoples (Clay, 1990). The Chechens, for example, lost at least one-quarter and perhaps half of their population in transit when they were deported *en masse* to Kazakhstan and Siberia in 1944. In 1957 they were allowed to return to their ancestral territory. In the face of continued Chechen rebellion to Russian appropriation of their land, economic resources, and a continued denial of civil rights, in late 1992

Russia sent tanks and troops to the north Caucasus, ostensibly as peacekeepers in an ethnic dispute.

While not all states are actively seeking the eradication of minorities within their borders, they have pursued policies designed to assimilate minorities into the mainstream or dominant culture. It was not too long ago that minority children in places such Australia, the United States, Britain, and Scandinavia were subject to physical violence in school for speaking their home language. Often the education of these children entailed removing them from their parents and their own cultural group. The Statutes of Iona in Scotland, dating from 1609, are among the early instances of legislation in present day U.K. designed to promote linguistic and cultural assimilation. The statutes had the expressed purpose of separating Highland children from their native Gaelic culture and language and educating them in English in the Lowlands, where they would not only learn the dominant language, but would do so in an alien cultural and linguistic environment where their own culture was seen as barbaric. The law required the chiefs to send their eldest child to the Lowlands to be educated until they could speak, read, and write English.

In North America native children were sent to boarding schools where their own languages were forbidden. In Canada, the federal government and churches entered into a formal partnership to run a residential school system for Indian and Inuit children as part of the assimilation policy of the Canadian government. Education in such church-run, government-funded residential schools was supposed to prepare children for life in white society by denying them their native identity. The residential school system was in operation for nearly 150 years. In some parts of Canada as many as five generations of children attended, and some communities were depopulated of children between the ages of 5 and 20. Such schooling produced a collective sense of shame about native languages and identities. It is not surprising that demands for some form of bilingual education emerge when a group feels it is being discriminated against on other grounds. In a study done of 46 linguistic minorities in 14 European countries, the clearest link to emerge between language and schooling is that a minority language which is not taught tends to decline (Allardt, 1979).

The borders of most countries are often linguistically diverse areas. Due to a variety of political and historical factors, bilinguals may be concentrated in particular geographic areas constituting regions where the use of a language other than the state language is normal. The northeastern corner of Italy shares a border with Slovenia to the east and Austria to the north. It contains a substantial population speaking either Slovenian, as well as Friulian (more closely related to Provençal than to standard Italian), or German (Standard German). Sauris is in effect a German linguistic island severed from the Austrian empire and incorporated into the Italian state. The region of Trentino-Alto Adige (Südtirol), governed by a special statute giving equal status to German (Standard German) and Italian, guarantees the right to education in the mother tongue for Germans in the province (from nursery to higher level). Italian is taught as a second language starting from the second year of the elementary cycle. Friulian is one of the largest minority languages of Italy, with over half a million speakers in the region of Friuli-Venezia-Giulia. A regional act of April 1993 provided funds for the promotion of Friulian in primary schools. Friulian is also used in some bilingual preschool education in the province of Udine. In the south and on the east coast, Greek and Albanian are spoken in some communities by descendants of refugees and mercenaries. Neither Greek nor Albanian has any official status, although the languages are taught in a small number of schools.

Although Article 6 of the Italian constitution is a clause pertaining to linguistic minorities which states that the republic protects linguistic minorities by special laws, there are discrepancies between policy and practice. Many minorities do not benefit from any special provisions. There are approximately one million speakers of Sardinian, which has no official recognition, despite the fact that Sardinia is an autonomous region governed by special statutes. Sardinian may be used in preprimary schools if needed to communicate with children. At the primary and secondary levels Sardinian has recently been introduced as a separate subject on an experimental basis.

In some countries decisions about language policy follow a 'personality' or 'territory' principle. In Switzerland territorial unilingualism exists under federal multilingualism in the country's four officially declared national languages: German, French, Italian, and Romansch. Of the 26 cantons, 22 are officially monolingual, with one of the four languages functioning as the dominant language in education. English is much preferred over the other official languages as a second language learned at school. Canada follows the personality principle for its two official languages, French and English, where sufficient numbers warrant. Quebec gives a universal right to French education, but the right to English education is limited to those with at least one parent educated in English.

Language Policy beyond the Nation-State

Many nation-states are similar to Italy in their incorporation of a number of groups with distinct languages, and their recognition of only one or a few languages for use within the education system and for other societal institutions. As the official language of Italy, Italian is also recognized beyond its national borders as an official language of the European Union (EU). Europe is perhaps unique in having such a large concentration of world languages within its borders as well as a sizable number of minority languages. From its beginning as the European Economic Community, the EU has accorded official status to each of the national languages of its member states. This means that relatively small national languages such as Danish, with roughly five million speakers, and Greek are, in principle, as official community languages on an equal footing with international languages such as French and English, a status they have nowhere else in the world. Outside the EU and its own borders, Danish has a similar status only in the Nordic Parliament, and, like Greek, it is not spoken at international gatherings. As embodied in its linguistic policy, this has meant that equality of access to the EU's institutions should not be hindered by language. In 1990 the European Parliament adopted its so-called 'principle of complete multilingualism,' which it declared to be "consistent with the respect which is owed to the dignity of all languages which reflect and express the cultures of the different peoples who make up the European community." However, from the beginning not all languages have been equal; nor in the larger sense was or could multilingualism ever be 'complete.' This resolution was adopted due to pressure in support of granting Catalan some sort of official standing in the EU's operations.

The case of Catalan is indicative of the fact that many minority languages, both indigenous and non-indigenous, are not recognized either as official or as working languages, even though some of them have larger numbers of speakers than do the national languages. Thus, Catalan with its roughly six million speakers, despite having more speakers than Danish, was not an official language because the country in which it was officially recognized, Andorra, is not a member of the EU. In the member states where it is spoken, France, Spain and Italy, it does not have official status. While denying official status to some languages like Catalan, the regulations of the EU have continually been expanded to accommodate the entrance of new member states with their national languages. In 2004 the EU expanded from 15 countries with 11 official languages to 25 countries with 20 official languages.

The EU has undertaken legislation to defend the status of certain minority language communities within its borders in the form of the European Charter for Regional and Minority Languages (1992). Although it specifies no list of actual languages, the languages concerned must belong to the European cultural tradition (which excludes 'immigrant' languages), have a territorial base (which excludes languages such as Yiddish [Western Yiddish] and Romany [Romani], used over a wide geographic area), and be a separate language identifiable as such (which excludes local dialects of the official or majority languages). The terms of reference are deliberately vague in order to leave open to each member state how to define cultural heritage and territory. The charter provides a large number of different actions that state parties can take to protect and promote historical regional and minority languages, from which states must agree to undertake at least 35. However, each state is free to name the languages which it accepts as being within the scope of the charter (Ó Riagáin, 1998). The U.K., for instance, ratified the treaty in March 2001, but did not include Manx and Cornish. Mercator Education maintains a database of information relating to the use of regional and minority languages in education.

Language planning on an even more limited regional basis clearly makes better sense for languages such as Saami, Basque, Catalan (Catalan-Valencian-Balear), and other languages cutting across national boundaries, but the EU has generally avoided taking any action which would interfere with national laws or policies concerning linguistic minorities. The result is that many languages are valued only beyond their national borders, while not being recognized for educational or other public purposes even within their own areas of concentration.

Language Policies at the International Level

The issue of language in education has been central to the mandate of UNESCO because one of its goals is to achieve universal quality primary education and a 50% increase in adult literacy by the year 2015. Established in 1945 as a special agency of the United Nations, the organization promotes international cooperation among its 190 member states and six associated members in the fields of education, science and culture. It aims to be a standard setter in forging international agreements on a variety of ethical issues. In 1953 UNESCO published an expert report

on the use of vernacular languages in education, whose recommendations are still considered to be a central reference and have been widely referred to.

Nevertheless, UNESCO's (1935: 6) much-cited axiom "that the best medium for teaching is the mother tongue of the pupil" did not lead to any widespread adoption and development of vernacular languages as media of education. Despite some encouraging developments in some countries, in most parts of the world schooling is still virtually synonymous with learning a second language. Brenzinger (1998) estimated that fewer than 10% of African languages are included in bilingual education programs, with the result that more than 1000 African languages receive no consideration in the education sector. Skutnabb-Kangas (2000) maintained that education for minorities in many parts of the world still operates in ways that contradict best practices. She estimated that fewer than 10% of the world's languages are used in education.

In 2003 UNESCO published a new position paper on languages and education reflecting the changing global context for education in a multilingual world (*see* **Education in a Multilingual Society**). The recommendations included choice of the language of instruction in multilingual contexts; the need to preserve the languages and the ethnic identities of small language groups; and the role of English as the lingua franca and the language of instruction in countries where it is not a native language. These concerns grew out of recognition of education as an important tool and reflection of cultural diversity in a rapidly changing world.

The changing character of multilingualism in the world today has manifested itself in at least two patterns. The first is that over the last few centuries in particular, some languages have shown a remarkable propensity to spread (*see* **Language Spread**). Speakers of the 10 largest languages make up about half the world's population, and this figure is increasing. The 100 largest languages account for 90% of all people, with the remaining 6000-some confined to ten percent of the world's most marginalized peoples, who have generally been on the retreat for several hundred years (*see* **Endangered Languages**). European colonization of the New World created many such language spreads, and most of the largest European languages are also widely spoken outside Europe. Today an Indo-European language, either English, French, Spanish or Portuguese, is the dominant language and culture in every country in North, Central and South America (Nettle and Romaine, 2000).

A second noteworthy trend is increasing bilingualism in a metropolitan language, particularly English, which has become the language of the 'global village.' No one knows exactly how many people speak English as a first or second language, but some estimates for the former group are 375 million; for the latter group, some figures run as high as 1.5 billion (roughly a quarter of the world's population). As the world's economy has shifted from an industrial base to one based on exchange of information, the globalizing new world order is founded on communications technology, which underlies the linking of national economies. Hence the role of language and communication is destined to play a more critical role than ever before (*see* **Languages of Wider Communication**). Because the technology facilitating these developments originated largely in the English-speaking world, English is at the leading edge of global scientific and economic development. As much as 80% of the information stored in the world's computers is in English and 90% of the world's computers connected to the Internet are located in English-speaking countries. English is now the most widely used language in publication, with over 28% of the world's books printed in English and over 60 countries publishing books in English. English is also the language of international air traffic control and the basis for Seaspeak, used in international maritime communication. Crystal (1997) estimated that 85% of international organizations use English as one of their working languages, among them the United Nations and its subsidiary organs. French is the only real rival to English in this arena and it has been continually losing ground. Virtually all major corporations advertise their products in English. English is also the language of international popular culture for today's youth.

Most people in northern European countries such as the Netherlands, Germany, and the Scandinavian countries are becoming bilingual in English at an increasingly earlier age through schooling. Soon there will be few monolinguals among their school-age populations. English has rapidly become the first preferred foreign language study at school in the European Union, with nearly 90% of students studying it. French is almost always the second most widely taught language. The teaching of one or more foreign languages in primary school has also become more widespread.

Many countries have changed their educational practices regarding the teaching of foreign languages as a response to increasing demand for English. In Iceland, for instance, English has replaced Danish as the first foreign language taught in compulsory education (i.e., primary and lower secondary) in the new national curriculum. Danish is still taught as a compulsory second language to maintain and strengthen ties and cooperation with other Nordic

countries. English instruction begins at age 10 (the fifth year of schooling) and is taught for 6 years, while Danish begins at age 12 and is taught for 4 years. During the last two years of lower secondary schooling students generally have the option of learning a third foreign language, usually German (Standard German), but Spanish and French in some cases. The national curriculum guidelines also prescribe a minimum number of hours per week of foreign language instruction. These are rather low: 16 hours per week in English over a 6 year period, 14 for Danish, and only 2 hours per week in the optional third foreign language. Foreign languages are generally not used as the medium of instruction.

In other parts of the world English has rapidly replaced other languages once widely taught as second languages. Under the Soviet regime Russian was imposed in schools throughout the former Soviet bloc. After disintegration of the Soviet Union, few countries besides Russia require students to learn it, with the result that the language is less and less used.

Meanwhile, a third trend is that immigration and migration have brought about increasing linguistic and cultural diversity in much of Europe as well as the United States and other parts of the globe (*see* **Migration and Language**). At the end of the 20th century one-third of the urban population in Europe under the age of 35 was composed of ethnic minorities, the result of widespread migration in the 1950s and 1960s when Europe experienced an acute labor shortage. Around 10% of the school age population already has a culture and language different from that of the majority of the country in which they reside (Extra and Verhoeven, 1999). London has become an increasingly diverse city with as many as 200 languages spoken in its schools as a result of the influx of overseas migrants from the Caribbean and Asia. Similarly, Melbourne, once primarily a monolingual city, now has the largest concentration of Greek speakers in the world.

At the beginning of the 20th century one in eight persons in the United States was nonwhite; by the end of the century the proportion had increased to one in four. The white population also grew more slowly than any other group in the latter half of the 20th century. From 1980 to 2000 the Hispanic population in the United States doubled. The U.S. Census 2000 revealed that persons claiming Hispanic or Latino origin have replaced African-Americans as the largest ethnic minority group. A third of California's population belongs to this minority and nearly 40% of its population claims to speak a language other than English at home (Hobbs and Stoops, 2002). The United States is now the fifth largest Hispanic country in the world. Cities such as Miami and Los Angeles are now predominantly Hispanophone, and Los Angeles has been Latinized by its continuing immigration from Mexico. In three states, California, New Mexico, and Hawaii, as well as the District of Columbia, minority populations constitute the majority.

Recognizing that issues of identity, power, and nationhood are closely linked to the use of specific languages in the classroom, UNESCO's (2003) position paper on education in multilingual contexts reaffirmed the value of mother tongues, but at the same time stressed the importance of balancing the need for local languages in learning and access to global languages through education. As far as mother tongue teaching is concerned, UNESCO advises that it should cover both teaching of and through this language for as long as possible (*see* **Mother Tongue Education: Standard Language**). Learning through a language other than one's own presents a double burden. Not only must new knowledge be mastered, but another language as well. Many minorities may be disadvantaged to begin with, coming from at-risk populations such as new immigrants, refugees, etc. (*see* **Immigrant Language Education**).

Linguistic Human Rights

The UNESCO (2003) position paper also endorsed many of the recommendations that have come out of the debate about linguistic human rights, which has emerged as an important topic in the context of the education of linguistic minorities (Varennes, 1996; Paulston, 1997; Skutnabb-Kangas, 2000). The notion of linguistic human rights is an attempt to link the debate about language rights with the relatively well-defined international legal framework in existence for human rights. That is, the concept of human rights is invoked as a means of reaching consensus on the rights of linguistic minorities to ensure social justice. These include the rights of indigenous and minority groups to education in their own language, access to the language of the larger community and that of the national education system, and international languages (see King and Schielmann, 2004). Discussion of a universal declaration of linguistic rights is taking place under the auspices of UNESCO. Such legislation aims at guaranteeing at an individual level that everybody can identify with their mother tongue(s) and have this identification accepted and respected by others, and can learn the mother tongue(s) fully, orally (when physiologically possible) and in writing. In most cases, this requires that indigenous and minority children be educated through the medium of their mother tongue(s); that they can use the mother tongue(s) in official situations (including schools);

that everybody whose mother tongue is not an official language in the country where they are resident can become bilingual (or multilingual, if they have more than one mother tongue) in the mother tongue(s) and (one of) the official language(s) (according to their own choice).

In practice, this is being achieved to some degree in some contexts, often by means of what has been called a 'three-language formula' or a '3 ± 1 language formula' in education (*see* **Education in a Multilingual Society**). In India a three-language policy means that children from non-Hindi-speaking areas study their regional language, in addition to Hindi, and English. Hindi speakers, on the other hand, study Hindi, English, and another language. In each state there is generally a large population who speak the dominant language of the neighboring state in addition to the dominant language of the state in which they reside. In Andhra Pradesh the dominant language is Telegu (Telugu), but many people speak Kannada, Marathi, and Tamil. Millions of Telegu speakers reside in the states of Karnataka, Orissa, Tamil Nadu and Maharashtra. Each state is multilingual and the linguistic majority in one state may be a minority in other states. Each state usually recognizes one official language and restricts the use of other languages to particular districts within the state. Critics of the policy contend that although it sounds fine in theory, in practice it has not been followed throughout the country.

In Luxembourg trilingualism in the national language, Luxembourgish or Lëtzebuergesch (Luxembourgeois), French, and German (spoken in the neighboring countries of Belgium, France, and Germany) is encoded in legislation which ensures that all citizens learn all three languages at school. Students begin with their everyday spoken language, Luxembourgish, in compulsory preschool education. German is added in the first year of primary education, and French from the second year of primary school onward. Over the years, however, and particularly in secondary education, French gets an ever-bigger share until it completely replaces German as the language of instruction. English is learned as a fourth compulsory language in secondary education and secondary technical education.

Kymlicka and Patten (2003: 7) claimed that one reason why there has been a general reluctance to view policies of official bilingualism as rights rather than as pragmatic accommodations is that public institutions in the most powerful Western nations, the U.K., the United States, France, and Germany, have been monolingual for a century or more with no significant movement toward challenging the hegemonic position of the majority language. Immigrants have not generally challenged the hegemony of these nations and have usually assimilated rapidly and none of these countries has faced the linguistic challenges of Belgium, Spain, Canada, or Switzerland. Language occupies a contested position when nation-states cannot ground their basis for a common identity on language, religion, or culture. Some still regard the concept of language rights as 'regressive' because they are seen as encouraging the persistence of ethnic differences, leading to conflict and divided loyalties. It is an unresolved question whether and when language shift can be required or expected in deliberative democracies (*see* **Language Maintenance and Shift**). Likewise, one can question whether it is legitimate for the state to insist that all children be schooled in the majority language of the state as the sole or main medium of instruction. National ethnic minorities have many more internationally and nationally coded rights than immigrants. The linguistic human rights movement has focused on securing a universal right to mother tongue primary education.

Typologies and Models of Multilingual Education

Bilingual education is not a modern phenomenon; it has existed in one form or another for at least 5000 years. Only recently, however, has it become an area of concern for policy makers. The term 'bilingual education' can mean different things in different contexts. If we take a commonsense approach and define it as a program where two languages are used equally as media of instruction, many so-called bilingual education programs would not count as such (*see* **Bilingual Education**). Moreover, the 'same' educational policy can lead to different outcomes, depending on differences in the input variables.

Typologies of bilingual education range from those which distinguish two basic types (Edwards, 1984) to Mackey's (1972) 90-cell typology. In her discussion of these many typologies each with somewhat different terminologies, Hornberger (1991) showed how the same terms are often confusingly used for different types of educational programs and conversely, different terms refer to the same type. So-called transitional bilingual education, for example, is also referred to as compensatory or assimilation bilingualism. Sometimes a distinction is made between immersion and submersion, and often the additional term 'structured immersion' is used for a program that has more in common with submersion than immersion. Like submersion, it is a program of monolingual majority language instruction for minority language speakers with little or no use of the pupils' first language. A so-called maintenance program does not necessarily

foster maintenance. Sometimes the term refers to a program's goal, e.g., the maintenance of a minority language, while in other cases, it refers to the structure of a program, e.g., the curricular maintenance of a minority language as a medium of instruction.

Hornberger proposed her own framework, which distinguishes between bilingual education models and program types. Models are defined in terms of their goals with respect to language, culture and society, and program types in terms of characteristics relating to student population, teachers, and program structure. This led her to recognize three types of models, transitional, maintenance, and enrichment, each of which may be implemented via a wide range of program types. Like Hornberger, Skutnabb-Kangas (2000) recognized three general types: immersion, submersion, and maintenance. If the educational aim of a bilingual program is the enrichment of majority children, an immersion program is chosen and the children are taught through the medium of a second language. The type of program chosen will typically, though not always, have different consequences. In practice, the situation in individual countries is complex and often several different options are available for different kinds of children, depending on a variety of circumstances, varying from place to place.

Immersion

Immersion programs first began in Montreal in 1965 to teach French to English-speaking students. Although there are many variants of the model, in most cases the students come from the same home language background, and the curriculum typically involves two or more languages as the medium of instruction. One of these is usually the student's home language, and the other a second or foreign language, with at least 50% of the curriculum being taught through the second or foreign language. The Canadian model can be thought of as leading to 'additive bilingualism' because the aim is to produce a high level of proficiency in both languages. Results have shown that immersion students consistently display normal levels of academic development in their first language while acquiring high levels of proficiency in the second language (*see* **Bilingualism and Second Language Learning**). Full benefits of immersion emerge after about 5 or 6 years of continuous participation. Early immersion programs tend to achieve better results, but late immersion programs can also be successful (Johnson and Swain, 1997; Cenoz and Genesee, 1998).

After the success of the French immersion programs, similar immersion models of various types have become widely used around the world to promote indigenous and minority languages (*see* **Language Education for Language Revival**). Some programs are total immersion, such as the Hawaiian program, which uses Hawaiian as the language across the curriculum. English is introduced as a subject from the fifth grade (around age 10) for 1 hour a day. Most of the students attending are English speakers and are learning Hawaiian as a second language. The immersion model contrasts with more conventional language teaching as a subject for a limited number of hours with fewer opportunities for high levels of academic or informal engagement with the language in use. In immersion there may be little if any focus on language learning *per se* in the form of direct teaching of grammar and vocabulary. Language is acquired through the meaningful interaction required to learn academic content in various subjects. Other variants of the model may rely on bilingual immersion combined with a third language taught as a subject. In parts of the Basque country Basque and Spanish are used for instruction during primary education, and English is taught as a subject beginning in kindergarten.

In other cases, however, total immersion programs have been used as a tool to assimilate linguistic minorities into dominant languages. Minority children are put into majority language classes (with or without some additional teaching of the second language). Some researchers have called such programs 'submersion' or 'subtractive bilingualism' since the second language gradually undermines proficiency in the first because the development of the child's first language is disrupted and incomplete. In the United States some children have received assimilationist treatment (with or without special instruction in English as a second language), while others have had the opportunity to participate in bilingual programs along with majority children who were being exposed to the first language of the minority group in an enrichment scheme. The types of programs offered to particular groups depend very much on the relationship between them and the government.

Transitional Bilingual Education

Bilingual education was given a legal footing in the United States by the Bilingual Education Act of 1968. Aimed at children with 'limited proficiency in English,' it provided funds for instruction in the mother tongue only as an aid to allow the children to proceed as rapidly as possible into ordinary mainstream classes in the majority language. From the beginning there was conflict over the degree of emphasis to be given to native language instruction.

The model of bilingual education prescribed by the federal government, however, was opposed in its aim and principles to the kind of enrichment goals underlying Canadian immersion. Although it provided opportunities for schools to set up bilingual education programs, it did not place individual schools under any legal obligation to do so. Moreover, there was no intention or provision to maintain the students' home language. Instead of receiving equal instruction in both languages as they would in a maintenance program, the students would be given increasingly less instruction in their native language until they finally left the program.

Litigation brought to the courts on behalf of various groups of minority students led in some cases to court-mandated bilingual education programs. In Lau vs. Nichols a class action suit was brought against the San Francisco Unified School District by Chinese public school students in 1970. It was argued that no special programs were available to meet the linguistic needs of these students. As a consequence, they were prevented from deriving benefit from instruction in English and were not receiving equal treatment. The plaintiffs made their appeal not on linguistic grounds, but on the basis of the Civil Rights Act of 1964, which states that "no person in the United States shall, on the ground of race, color or national origin, be excluded from participation in, be denied the benefits of, or be subject to discrimination under any program or activity receiving Federal financial assistance" (Teitelbaum and Hiller, 1977: 6). In their case against the school board, the plaintiffs requested a program of bilingual education. Although the case was lost, the Supreme Court overturned the decision of the federal district court in 1974. It concluded that "the Chinese-speaking minority receives fewer benefits than the English-speaking majority from respondents' school system which denies them a meaningful opportunity to participate in the educational program – all earmarks of discrimination banned by the regulations" (Teitelbaum and Hiller, 1977: 8). This was a landmark decision because it meant that for the first time in the United States the language rights of non-English speakers were recognized as a civil right. It was one of the few language cases ever to reach the Supreme Court, and its ruling made schools rather than parents or children responsible for remedying the children's limited knowledge of English.

In its decision the Supreme Court did not press for any specific remedy. It pointed out only two possibilities: namely, teaching English to the students or teaching them in Chinese. They requested only that the school board rectify the situation of inequality of educational opportunity. The remedy taken by the San Francisco school board was to set up a bilingual education program for Chinese, Filipino, and Spanish language groups, who made up over 80% of the students with little or no English. Teaching in English as a second language was offered to all other minority groups. The Lau decision led to other cases. It also encouraged expansion of the services and eligibility provided through the Bilingual Education Act, and many states passed bills mandating bilingual education. The Lau decision was also instrumental in setting up policy guidelines at the federal level that would allow the U.S. Office of Education to decide whether a school district was in compliance with the Civil Rights Act and the Lau case. A document referred to as the 'Lau Remedies' directed school boards to identify students with a primary or home language other than English and to assess their proficiency in English and the home language. Elementary school students were to be taught in their dominant language until they were able to benefit from instruction entirely in English. The U.S. Congress made the Lau decision an explicit part of the Equal Educational Opportunities Act (1974). Further mandates for bilingual instruction followed from lawsuits by Latino parents.

The sixth (and final) version of the Bilingual Education Act reauthorized by Congress in 1994 endorsed for the first time the goal of developing native language skills alongside its traditional focus on English language acquisition for limited-English-proficient children. Meanwhile, under the Clinton administration (1993–2001) two-way or dual immersion programs were promoted; these grew more than tenfold between 1987 and 2001. These were aimed at integrating majority and minority children in a program of content and literacy instruction in two languages.

In the late 1990s, however, the tables turned dramatically when a lawsuit was filed on behalf of Latino parents who claimed that state policies mandating Spanish instruction discriminated against their children (Carbajal *et al*. vs. Albuquerque Public Schools, 1999). This case attempted to portray bilingual education as a violation of civil rights rather than an entitlement. In California a conservative software millionaire named Ron Unz spearheaded a movement named English for the Children, portraying itself as a group committed to securing the right to instruction in English for immigrants. Under this proposal children with limited English proficiency were to be 'mainstreamed' as soon as possible into regular classrooms. After voters in three states (California in 1998, Arizona in 2000, and Massachusetts in 2002) voted against bilingual education, programs were dismantled, and have been under

attack in others. This outcome meant that most of the bilingual education programs enrolling 43% of the English language learners in the United States would be replaced with intensive English immersion.

Ironically, one of the reasons why many people have viewed bilingual education so negatively is due to the fear that it aims to maintain languages, and by implication cultures, other than English. Even at the peak of their existence, bilingual programs reached only a minority of children for whom they could have been beneficial. Voters appeared to be largely ignorant of the rationale behind such programs as well as their aims and outcomes. Crawford's (2004) analysis suggested that the public had a mistaken view that bilingual education was a diversion from acquiring English rather than a means to that end. It was the idea of maintaining languages other than English which the public was against. While foreign language instruction in the world's major languages in mainstream schools has been seen as valuable, both economically and culturally, bilingual education for minority students has been equated with poverty and loyalties to nonmainstream culture which threaten the cohesiveness of the state. Voters were also misled by the use of the term 'English immersion,' which suggested an intensive English program tailored to the needs of children learning English.

In addition, various opponents of bilingual education have formed a powerful lobby backed by considerable sums of money. The English-only movement formed in 1983, operating under the name U.S. English, has been campaigning in favor of a constitutional amendment to make English the official language of the United States, and for similar legislation at state level. The organization also seeks to repeal laws mandating multilingual ballots and voting materials. Twenty-seven states have enacted some form of official legislation. The group has seen programs that accommodate immigrants in their native languages as a kind of 'linguistic welfare' system that lowers the incentive to learn English and restricts them to low-skilled, low-paying jobs.

In 2002 the Congress effectively repealed the Bilingual Education Act when it passed the 'No Child Left Behind Act.' References to bilingualism have been removed from various federal agencies to reflect the shift away from bilingual education to concentration on the acquisition of English. The Office of Bilingual and Minority Affairs at the U.S. Department of Education was renamed the Office of English Language Acquisition, Language Enhancement and Academic Achievement for Limited English Proficient Students. The National Clearing House for Bilingual Education has been renamed the National Clearing House for English Language Acquisition and Language Instruction Education Programs.

As in the case of immersion programs aimed at enrichment, research supports the pedagogic effectiveness of bilingual education programs. Students in bilingual programs have typically performed at least as well on English reading tests as students in all-English programs (Corson, 2001; Crawford, 2004). Much of the concern over acquisition of English masks the fear many middle-class whites have of losing their majority status. Although proponents of U.S. English attempt to legitimize the organization's existence as a way of breaking down supposed language barriers and facilitating minority access to the material and other benefits of mainstream America, the irony is that most ethnic minorities do not actually want a self-contained ethnic group where no English would be spoken. Nor, however, do they want to assimilate linguistically or culturally. A majority want to maintain their ethnicity and language while also being American. An unfounded fear of diversity itself and thinly disguised racism lies behind the backlash against bilingual education in the United States, and in Europe, where earlier policies of providing home language instruction to the children of migrant workers have been rescinded or drastically curtailed.

Despite propaganda from U.S. English to the contrary, there were actually 4.5 times as many non-English speakers recorded in the U.S. census when immigration reached its highest level than in the 1990 census. The assimilative forces that absorbed those immigrants and their languages are even more powerful today. Although the number of non-English speakers is increasing, so too is the rate of shift to English. Languages other than English are the ones under threat. Spanish is fast approaching a two-generation pattern of language shift rather than the three-generation model typical of immigrant groups in the past. Without the replenishing effects of continuing immigration, Spanish would scarcely be viable in the United States over the long term (Veltman, 1983, 1988).

Weak Linkages between Language Policy and Planning

Despite evidence of growing rather than decreasing diversity in many education systems, in some countries the trend has been toward not recognition of the need for policy and planning but the imposition of ever more centralized provision and greater intolerance of diversity. There are important differences between 'tolerance rights' and 'promotion rights.'

Most democracies provide for freedom of government interference in private language use, but many are reluctant to make legal provision for promotion of languages in the public sector other than the dominant language(s).

In addition, weak linkages between policy and planning render many existing policies ineffective (Romaine, 2002). Many language policy statements are often reactive *ad hoc* declarations lacking a planning element. Eritrea's 1995 decision **not** to recognize an official language is a grand declaration with no linkage between policy and planning. Thus, President Isayas Afewerki (Brenzinger, 1998: 94):

> When we come to the question of language as a means of instruction in schools, our principle is that the child should use its mother tongue or a language chosen by its parents in the early years of its education, irrespective of the level of development of the language. Our policy is clear and we cannot enter into bargaining. Everyone is free to learn in the language he or she prefers, and no one is going to be coerced into using this or that 'official' language.

Policies cannot be implemented unless those with the duty to implement them are provided with the necessary resources. Policies cannot be effective unless they are tied to a plan for monitoring of compliance and application of sanctions where they are not implemented.

See also: Bilingual Education; Bilingualism and Second Language Learning; Bilingualism; Education in a Former Colonial Language; Education in a Multilingual Society; Education in a Multilingual Society; Educational Failure; Endangered Languages; Immigrant Language Education; Language Education for Language Revival; Language Maintenance and Shift; Language Planning and Policy: Models; Language Spread; Languages of Wider Communication; Law and Language: Overview; Migration and Language; Minorities and Language; Mother Tongue Education: Standard Language.

Bibliography

Allardt E (1979). *Implications of the ethnic revival in modern industrialized society: a comparative study of the linguistic minorities in western Europe.* Helsinki: Societas Scientariarum Fennica.

Baker C & Prys Jones S (1998). *Encyclopedia of bilingualism and bilingual education.* Clevedon, England: Multilingual Matters.

Brenzinger M (1998). 'Various ways of dying and different kinds of death: scholarly approaches to language endangerment on the African continent.' In Matsumura K (ed.) *Studies in endangered languages: papers from the International Symposium on Endangered Languages, Tokyo, November 18–29*, 1995. Tokyo: Hituzi Syobo. 85–100.

Carbajal *et al.* v. Albuquerque Public Schools (1999). D. New. Mex. Case No. CIV 98–279 MV/DJS. May 11.

Cenoz J & Genesee F (1998). *Beyond bilingualism: multilingualism and multilingual education.* Clevedon, England: Multilingual Matters.

Clay J (1990). 'Indigenous peoples: the miner's canary for the twentieth century.' In Head S & Heinzman R (eds.) *Lessons of the rainforest.* San Francisco: Sierra Club Books. 106–117.

Corson D (1990). *Language policy across the curriculum.* Clevedon, England: Multilingual Matters.

Corson D (2001). *Language diversity and education.* Mahwah, NJ: Lawrence Erlbaum Associates.

Crawford J (2004). *Educating English learners: language diversity in the classroom.* Los Angeles: Bilingual Education Services.

Crystal D (1997). *English as a global language.* Cambridge: Cambridge University Press.

Edwards J R (ed.) (1984). *Linguistic minorities, policies and pluralism.* London & New York: Academic Press.

Extra G & Verhoeven L (eds.) (1993). *Immigrant languages in Europe.* Clevedon, England: Multilingual Matters.

Extra G & Verhoeven L (eds.) (1999). *Bilingualism and migration.* Berlin: Mouton de Gruyter.

Fishman J A, Ferguson C A & Das Gupta J (eds.) (1968). *Language problems of developing nations.* New York: John Wiley & Sons.

Grimes B F & Grimes J E (eds.) (2000). *Ethnologue: languages of the world,* (14th ed.). Dallas: SIL International.

Hobbs F & Stoops N (2002). *Demographic trends in the 20th century.* Washington, DC: U.S. Government Printing Office.

Hornberger N H (1991). 'Extending enrichment bilingual education: revisiting typologies and redirecting policy.' In García O (ed.) (1991) *Bilingual education: Focusschrift in honor of Joshua A. Fishman.* Amsterdam & Philadelphia: John Benjamins. 215–234.

Johnson K & Swain M (1997). *Immersion education: international perspectives.* Cambridge: Cambridge University Press.

King L & Schielmann S (eds.) (2004). *The challenge of indigenous education: practice and perspectives.* Paris: UNESCO.

Koenig M (ed.) (2001). *MOST Journal on Multicultural Societies 3(2): The Human Rights of Linguistic Minorities and Language Policies.*

Kymlicka W & Patten A (eds.) (2003). *Language rights and political theory.* Oxford: Oxford University Press.

Lau v. Nichols. (1974). 414 U.S. 563. United States Congress.

Mackey W F (1972). 'A typology of bilingual education.' In Fishman J A (ed.) *Advances in the sociology of language.* The Hague: Mouton. 413–432.

May S (2001). *Language, education and minority rights: ethnicity, nationalism and the politics of language.* Harlow, England: Longman.

Nettle D & Romaine S (2000). *Vanishing voices: the extinction of the world's languages*. Oxford: Oxford University Press.
Ó Riagáin D (ed.) (1998). *Vade-Mecum: a guide to legal, political and other official international documents pertaining to the lesser used languages of Europe*. Dublin: European Bureau for Lesser Used Languages.
Paulston C B (1997). 'Language policies and language rights.' *Annual Review of Anthropology 26,* 73–85.
Romaine S (1995). *Bilingualism* (2nd ed.). Oxford: Blackwell.
Romaine S (2002). 'The impact of language policy on endangered languages.' *MOST Journal on Multicultural Societies 4(2),* 1–28.
Schiffman H (1996). *Linguistic culture and language policy.* London: Routledge.
Skutnabb-Kangas T (2000). *Linguistic Genocide in Education – or Worldwide Diversity and Human Rights?* Mahwah, NJ: Lawrence Erlbaum Associates.
Skutnabb-Kangas T & Phillipson R (eds.) (1994). *Linguistic Human Rights: overcoming linguistic discrimination.* Berlin & New York: Mouton de Gruyter.
Spolsky B (2004). *Language policy.* Cambridge: Cambridge University Press.
Teitelbaum H & Hiller R J (1977). 'The legal perspective.' In *Center for Applied Linguistics, Bilingual education: current perspectives, vol. 3: Law.* Arlington, VA: Center for Applied Linguistics. 1–67.
Tollefson J (ed.) (1995). *Power and inequality in language education.* Cambridge: Cambridge University Press.
Tollefson J (ed.) (2001). *Language policies in education: critical issues.* Mahwah, NJ: Lawrence Erlbaum Associates.
UNESCO (1953). *The use of vernacular languages in education.* Paris: UNESCO.
UNESCO (2003). *Education in a multilingual world.* Paris: UNESCO.
Varennes F de (1996). *Language, minorities and human rights.* The Hague: Martinus Nijhoff.
Varennes F de (2001). *A guide to the rights of minorities and language.* Budapest: Constitutional and Legal Policy Initiative.
Veltman C J (1983). *Language shift in the United States.* Berlin: Mouton.
Veltman C J (1988). *The future of the Spanish language in the United States.* Washington, DC: Hispanic Development Project.

Relevant Websites

http://www.ethnologue.com – Ethnologue.
http://www.mercator-education.org – Mercator Education.
http://www.unesco.org – MOST Clearing House.
http://www.us-english.org – U.S. English.

Language Politics

U Ammon, University Duisburg-Essen, Duisburg, Germany

Language politics (LPt) is actual political activity with respect to language and can be distinguished from *language policy* (LPc), i.e., the conception and planning of such activity; but both are, as a rule, closely intertwined (Spolsky, 2004; Ozolins, 1996). LPt's typical agents are state politicians, but they can be members of any polity (institution or community) – for example, a family, a religious community, a town, a nongovernmental organization, or an international organization. Objectives of LPt can vary greatly depending on interests and motives – e.g., 'purifying' one's own national language of foreign loans to shape it into a more adequate symbol of national identity, or spreading the language within the state or beyond in order to more efficiently exert power. LPt has to reckon with existing language rights and may result in new language rights (Paulston, 1997) (*see* **Linguistic Rights**; **Linguistic Decolonization**).

Internal LPt regulates language within the polity, while *external* LPt aims beyond it. For the state, the former is part of interior politics, the latter of exterior politics. Internal LPt can be directed at language structure (*language corpus politics*, LCPt) or language status and function (*language status politics*, LSPt). Typical aims of LCPt are graphization (introducing or regulating script and orthography); standardization, including codification (selection and codification of norms of spelling, pronunciation, vocabulary, grammar, and style or texts); 'purification' (eliminating foreign loans); and modernization (developing modern terminology). Typical aims of LSPt are spreading the norm of a standard variety and, in the case of multilingual communities, allocating languages to certain regions (McRae, 1975) or domains and functions – for example, official (Laitin, 1992), educational (medium or subject of teaching on various educational levels; Lo Bianco, 1987), religious (Landau and Kellner-Heinkele, 2001), the media, or the military (Ager, 1996; Schiffman, 1996). Politics of language promotion, language maintenance, or language revival can comprise

Nettle D & Romaine S (2000). *Vanishing voices: the extinction of the world's languages.* Oxford: Oxford University Press.

Ó Riagáin D (ed.) (1998). *Vade-Mecum: a guide to legal, political and other official international documents pertaining to the lesser used languages of Europe.* Dublin: European Bureau for Lesser Used Languages.

Paulston C B (1997). 'Language policies and language rights.' *Annual Review of Anthropology 26,* 73–85.

Romaine S (1995). *Bilingualism* (2nd ed.). Oxford: Blackwell.

Romaine S (2002). 'The impact of language policy on endangered languages.' *MOST Journal on Multicultural Societies 4(2),* 1–28.

Schiffman H (1996). *Linguistic culture and language policy.* London: Routledge.

Skutnabb-Kangas T (2000). *Linguistic Genocide in Education – or Worldwide Diversity and Human Rights?* Mahwah, NJ: Lawrence Erlbaum Associates.

Skutnabb-Kangas T & Phillipson R (eds.) (1994). *Linguistic Human Rights: overcoming linguistic discrimination.* Berlin & New York: Mouton de Gruyter.

Spolsky B (2004). *Language policy.* Cambridge: Cambridge University Press.

Teitelbaum H & Hiller R J (1977). 'The legal perspective.' In *Center for Applied Linguistics, Bilingual education: current perspectives, vol. 3: Law.* Arlington, VA: Center for Applied Linguistics. 1–67.

Tollefson J (ed.) (1995). *Power and inequality in language education.* Cambridge: Cambridge University Press.

Tollefson J (ed.) (2001). *Language policies in education: critical issues.* Mahwah, NJ: Lawrence Erlbaum Associates.

UNESCO (1953). *The use of vernacular languages in education.* Paris: UNESCO.

UNESCO (2003). *Education in a multilingual world.* Paris: UNESCO.

Varennes F de (1996). *Language, minorities and human rights.* The Hague: Martinus Nijhoff.

Varennes F de (2001). *A guide to the rights of minorities and language.* Budapest: Constitutional and Legal Policy Initiative.

Veltman C J (1983). *Language shift in the United States.* Berlin: Mouton.

Veltman C J (1988). *The future of the Spanish language in the United States.* Washington, DC: Hispanic Development Project.

Relevant Websites

http://www.ethnologue.com – Ethnologue.
http://www.mercator-education.org – Mercator Education.
http://www.unesco.org – MOST Clearing House.
http://www.us-english.org – U.S. English.

Language Politics

U Ammon, University Duisburg-Essen, Duisburg, Germany

Language politics (LPt) is actual political activity with respect to language and can be distinguished from *language policy* (LPc), i.e., the conception and planning of such activity; but both are, as a rule, closely intertwined (Spolsky, 2004; Ozolins, 1996). LPt's typical agents are state politicians, but they can be members of any polity (institution or community) – for example, a family, a religious community, a town, a nongovernmental organization, or an international organization. Objectives of LPt can vary greatly depending on interests and motives – e.g., 'purifying' one's own national language of foreign loans to shape it into a more adequate symbol of national identity, or spreading the language within the state or beyond in order to more efficiently exert power. LPt has to reckon with existing language rights and may result in new language rights (Paulston, 1997) (*see* **Linguistic Rights**; **Linguistic Decolonization**).

Internal LPt regulates language within the polity, while *external* LPt aims beyond it. For the state, the former is part of interior politics, the latter of exterior politics. Internal LPt can be directed at language structure (*language corpus politics*, LCPt) or language status and function (*language status politics*, LSPt). Typical aims of LCPt are graphization (introducing or regulating script and orthography); standardization, including codification (selection and codification of norms of spelling, pronunciation, vocabulary, grammar, and style or texts); 'purification' (eliminating foreign loans); and modernization (developing modern terminology). Typical aims of LSPt are spreading the norm of a standard variety and, in the case of multilingual communities, allocating languages to certain regions (McRae, 1975) or domains and functions – for example, official (Laitin, 1992), educational (medium or subject of teaching on various educational levels; Lo Bianco, 1987), religious (Landau and Kellner-Heinkele, 2001), the media, or the military (Ager, 1996; Schiffman, 1996). Politics of language promotion, language maintenance, or language revival can comprise

LCPt, such as constructing or reconstructing vocabulary, as well as LSPt, such as encouraging use in the family or institutionalization in school (Ó Riagáin, 1997) (*see* **Applying Pragmatics**).

External LPt is mainly LSPt or, more specifically, *language spread politics*. Examples of global consequences occurred during colonialism (Phillipson, 1992) (*see* **Linguistic Imperialism**). Today virtually all the larger language communities, or their major states, endeavor to allocate their languages in international organizations (e.g., the Arab countries in the UN in 1973, France and Germany in the EU; Ammon, 2003, 2004) or in the school curricula of other countries (Ammon and Kleineidam, 1992, 1994; Ricento, 2001; Haarmann, 2005).

See also: Applying Pragmatics; Endangered Languages; Identity and Language; Language Maintenance and Shift; Language Planning and Policy: Models; Language Policies: Policies on Language in Europe; Language Policy in Multilingual Educational Contexts; Languages of Wider Communication; Law and Language: Overview; Law on Language; Linguistic Decolonization; Linguistic Rights; Media and Language: Overview; Politics and Language: Overview; Pragmatics: Linguistic Imperialism; Radio: Language; Reversing Language Shift; Society and Language: Overview; Standardization; Teaching of Minority Languages; Television: Language.

Bibliography

Ager D E (1996). *Language policy in Britain and France.* London and New York: Cassell.

Ammon U (2003). 'Sprachenpolitik in Europa – unter dem vorrangigen Aspekt von Deutsch als Fremdsprache.' *Deutsch als Fremdsprache 40,* 195–209.

Ammon U (2004). 'Sprachenpolitik in Europa – unter dem vorrangigen Aspekt von Deutsch als Fremdsprache.' *Deutsch als Fremdsprache 41,* 3–10.

Ammon U & Kleineidam H (eds.) (1992). 'Language spread policy.' *International Journal of the Sociology of Language 95.*

Ammon U & Kleineidam H (eds.) (1994). 'Language spread policy.' *International Journal of the Sociology of Language 107.*

Haarmann H (2005). 'The politics of language spread.' In Ammon U *et al.* (eds.) *Sociolinguistics/Soziolinguistik,* vol. 2. Berlin and New York: de Gruyter.

Laitin D D (1992). *Language repertoires and state construction in Africa.* Cambridge: University Press.

Landau J & Kellner-Heinkele B (2001). *Politics of language in the ex–Soviet Muslim states.* London: Hurst.

Lo Bianco J (1987). *National policy on languages.* Canberra: Australian Government Publishing Service.

McRae K D (1975). 'The principle of territoriality and the principle of personality in multilingual states.' *International Journal of the Sociology of Language 4,* 33–54.

Ó Riagáin P (1997). *Language policy and social reproduction: Ireland, 1893–1993.* Oxford: Clarendon.

Ozolins U (1996). 'Language policy and political reality.' *International Journal of the Sociology of Language 118,* 181–200.

Paulston C B (1997). 'Language policies and language rights.' *Annual Review of Anthropology 26,* 73–85.

Phillipson R (1992). *Linguistic imperialism.* Oxford: University Press.

Ricento T (ed.) (2001). *Ideology, politics and language policies: focus on English.* Amsterdam and Philadelphia: Benjamins.

Schiffman H E (1996). *Linguistic culture and language policy.* London and New York: Routledge.

Spolsky B (2004). *Language policy.* Cambridge: University Press.

Language Processing: Statistical Methods

C Brew, The Ohio State University, Columbus, OH, USA

It is not a new idea to analyze language by counting things and uncovering statistical regularities in the results. Ancient applications include code-breaking and literary analysis, while the late 19th and early 20th century brought statistical approaches to language-related design problems in telegraphy and stenography. Researchers in the fields of computational linguistics, language technology, and natural language processing can expect to encounter statistical methods and probabilistic thinking on a daily basis. The key enabling factors for natural language processing were the arrival of the digital computer and the availability of automatic data processing. There is nevertheless a body of striking empirical work that predates the ready availability of computers, including Mosteller and Wallace (1984) on authorship attribution in the *Federalist papers* and Kučera and Francis (1967), who pioneered the design and creation of large-scale data resources for language analysis.

The focus of this article is on the statistical methods that have been developed for processing language by computer, but it is worth beginning with a detour that addresses two well-known objections. One is

LCPt, such as constructing or reconstructing vocabulary, as well as LSPt, such as encouraging use in the family or institutionalization in school (Ó Riagáin, 1997) (*see* **Applying Pragmatics**).

External LPt is mainly LSPt or, more specifically, *language spread politics*. Examples of global consequences occurred during colonialism (Phillipson, 1992) (*see* **Linguistic Imperialism**). Today virtually all the larger language communities, or their major states, endeavor to allocate their languages in international organizations (e.g., the Arab countries in the UN in 1973, France and Germany in the EU; Ammon, 2003, 2004) or in the school curricula of other countries (Ammon and Kleineidam, 1992, 1994; Ricento, 2001; Haarmann, 2005).

See also: Applying Pragmatics; Endangered Languages; Identity and Language; Language Maintenance and Shift; Language Planning and Policy: Models; Language Policies: Policies on Language in Europe; Language Policy in Multilingual Educational Contexts; Languages of Wider Communication; Law and Language: Overview; Law on Language; Linguistic Decolonization; Linguistic Rights; Media and Language: Overview; Politics and Language: Overview; Pragmatics: Linguistic Imperialism; Radio: Language; Reversing Language Shift; Society and Language: Overview; Standardization; Teaching of Minority Languages; Television: Language.

Bibliography

Ager D E (1996). *Language policy in Britain and France.* London and New York: Cassell.

Ammon U (2003). 'Sprachenpolitik in Europa – unter dem vorrangigen Aspekt von Deutsch als Fremdsprache.' *Deutsch als Fremdsprache 40,* 195–209.

Ammon U (2004). 'Sprachenpolitik in Europa – unter dem vorrangigen Aspekt von Deutsch als Fremdsprache.' *Deutsch als Fremdsprache 41,* 3–10.

Ammon U & Kleineidam H (eds.) (1992). 'Language spread policy.' *International Journal of the Sociology of Language 95.*

Ammon U & Kleineidam H (eds.) (1994). 'Language spread policy.' *International Journal of the Sociology of Language 107.*

Haarmann H (2005). 'The politics of language spread.' In Ammon U *et al.* (eds.) *Sociolinguistics/Soziolinguistik,* vol. 2. Berlin and New York: de Gruyter.

Laitin D D (1992). *Language repertoires and state construction in Africa.* Cambridge: University Press.

Landau J & Kellner-Heinkele B (2001). *Politics of language in the ex–Soviet Muslim states.* London: Hurst.

Lo Bianco J (1987). *National policy on languages.* Canberra: Australian Government Publishing Service.

McRae K D (1975). 'The principle of territoriality and the principle of personality in multilingual states.' *International Journal of the Sociology of Language 4,* 33–54.

Ó Riagáin P (1997). *Language policy and social reproduction: Ireland, 1893–1993.* Oxford: Clarendon.

Ozolins U (1996). 'Language policy and political reality.' *International Journal of the Sociology of Language 118,* 181–200.

Paulston C B (1997). 'Language policies and language rights.' *Annual Review of Anthropology 26,* 73–85.

Phillipson R (1992). *Linguistic imperialism.* Oxford: University Press.

Ricento T (ed.) (2001). *Ideology, politics and language policies: focus on English.* Amsterdam and Philadelphia: Benjamins.

Schiffman H E (1996). *Linguistic culture and language policy.* London and New York: Routledge.

Spolsky B (2004). *Language policy.* Cambridge: University Press.

Language Processing: Statistical Methods

C Brew, The Ohio State University, Columbus, OH, USA

It is not a new idea to analyze language by counting things and uncovering statistical regularities in the results. Ancient applications include code-breaking and literary analysis, while the late 19th and early 20th century brought statistical approaches to language-related design problems in telegraphy and stenography. Researchers in the fields of computational linguistics, language technology, and natural language processing can expect to encounter statistical methods and probabilistic thinking on a daily basis. The key enabling factors for natural language processing were the arrival of the digital computer and the availability of automatic data processing. There is nevertheless a body of striking empirical work that predates the ready availability of computers, including Mosteller and Wallace (1984) on authorship attribution in the *Federalist papers* and Kučera and Francis (1967), who pioneered the design and creation of large-scale data resources for language analysis.

The focus of this article is on the statistical methods that have been developed for processing language by computer, but it is worth beginning with a detour that addresses two well-known objections. One is

Jonathan Swift's satirical account of stochastic text generation in the academy at Lagado.

> [...] bits of Wood were covered on every Square with Paper pasted on them, and on these Papers were written all the Words of their Language, in their several Moods, Tenses, and Declensions, but without any Order. The Professor then desired me to observe, for he was going to set his Engine at Work. The Pupils at his Command took each of them hold of an Iron Handle, whereof there were fourty fixed round the Edges of the Frame, and giving them a sudden turn, the whole Disposition of the Words was entirely changed. He then commanded six and thirty of the Lads to read the several Lines softly as they appeared upon the Frame; and where they found three or four Words together that might make part of a Sentence, they dictated to the four remaining Boys who were Scribes. This Work was repeated three or four Times, and at every turn the Engine was so contrived that the Words shifted into new Places, as the Square bits of Wood moved upside down. (Swift, *Gulliver's travels,* chapter 5)

Swift's intent is to present as self-evidently absurd the notion that anything of literary value could be produced by such a process. This is satire and makes no scientific claim, but language technologists may nevertheless find more of scientific interest than Swift would have expected. Exactly how are the "Moods, Tenses, and Declensions" encoded? Why does the machine have 40 iron handles rather than just one? Are the words produced always drawn from the same probability distribution, or does the machine have a number of distinct internal states, so that sometimes one word is more likely, sometimes another? In spite of the satirical intent, interesting scientific issues are raised.

A second objection is made explicit in Noam Chomsky's discussion of statistics, which is more measured and subtle than Swift's satire.

> (1) Colorless green ideas sleep furiously
> (2) Furiously sleep green ideas colorless
> ... It is fair to assume that neither sentence (1) nor sentence (2) (nor indeed any part of these sentences) has ever occurred in English discourse. Hence, in any statistical model for grammaticalness, these sentences will be ruled out on identical grounds. (Chomsky, 1957: 16)

Chomsky's goal in making this argument was to contrast the statistical approach with his own preferred methodology, which relied primarily on introspection about the grammatical acceptability of sentences found or imagined by the linguist.

Chomsky presumably reached his conclusion because he was operating with a rather narrower notion of 'possible statistical model' than do most statistical linguists. For example, it is perfectly possible to construct a statistical model that sees not individual words but broad word classes. Simplifying somewhat, such a model might assign probabilities to (1) and (2) by treating them as if they were instead (3) and (4), respectively.

> (3) Adjective Adjective Noun Verb Adverb
> (4) Adverb Verb Noun Adjective Adjective

Under this scenario it is very much to be expected that (4) will receive a lower probability than (3), because it is so rare to see pairs of adjectives occurring at the end of an English sentence. Chomsky was not wholly averse to statistics and explicitly acknowledged that statistical methods can and should be used to model the use of language, but he did deny their usefulness as tools for explicating syntactic structure. The cited passage, and similar passages elsewhere in Chomsky's work, are a large part of the reason why most theoretical linguists have tended to avoid statistics. In spite of the scientific intent, much of the impact of Chomsky's discussion comes from the satirical bite of its rhetoric, rather than from the technical correctness of the arguments presented.

More detail on the relationship between statistics and linguistics is given by Abney (1996), Pereira (2000), and Bod *et al.* (2004).

For purposes of language technology, it is moot whether theoretical linguists accept statistical methods, since it is very clear that these methods can contribute to the creation of useful and instructive linguistic artifacts that carry out such tasks as part-of-speech assignment and word-sense disambiguation. The balance of this article gives an introduction to some of the concepts and techniques used.

Probability

'Probability' is the standard means for rational reasoning about uncertainty. In probability theory, a key concept is the 'event.' Examples of events include these:

- A coin is tossed and comes down heads.
- The result of the spin of a roulette wheel is 19.
- Two dice are rolled, and the sum of the results is 7.
- Two contiguous words are selected from Conan Doyle's 'A case of identity' and they turn out to be *Sherlock* and *Holmes*, respectively.

We can describe the situations above in terms of an event x and its probability $P(x)$.

- x: A coin is tossed and comes down heads. $P(x) = 0.5$ for a fair coin that comes down heads half the time.

- x: The result of the spin of a roulette wheel is 19. $P(x) = 1/38$ for a fair roulette wheel with 38 slots.
- x: Two dice are rolled, and the sum of the results is 7. $P(x) = 6/36 = 1/6$ for fair dice.
- x: Two contiguous words are selected from Conan Doyle's 'A case of identity' and they turn out to be *Sherlock* and *Holmes*, respectively. $P(x) = 7/7051$ because there are 7051 pairs of two contiguous words in the story, and exactly 7 of them are *Sherlock Holmes*.

These probabilities are obtained either by counting possibilities (for the coin or the dice) or by counting occurrences of events (as in the Conan Doyle example). By definition, a fair coin is one that makes the two possible outcomes equally likely. A fair die is one that gives a probability of 1/6 to each of the six possible outcomes. Therefore there are 36 equally likely outcomes to a roll of two dice, of which exactly six (1 and 6; 2 and 5; 3 and 4; 4 and 3; 5 and 2; 6 and 1) add up to the required number. The probability calculus supports this kind of reasoning by explaining how the probability of the compound event of rolling 7 is related to those of the simpler components of that event. Just as logic is a method for reasoning about the truth or falsity of propositions, so the probability calculus is a method for reasoning about uncertainty. The rules of logical inference remain the same regardless of the nature of the premises. In the same way, the rules of probabilistic inference remain the same regardless of the origin of the initial probability estimates.

Random Variables and Bayes's Rule

We have already presented an informal argument that links the probability of the event 'two dice are thrown and the total is 7' to the probabilities of the different possible outcomes of rolling the dice. A more formal presentation would be the following. Define two random variables R, G (corresponding to a red and a green die, respectively). Both R and G can take on any of the values 1, 2, 3, 4, 5, 6. We notate the probability of rolling a 1 with the green die as $P(1)$ and the probability of rolling a 3 with the green die and a 4 with the red die as $P(3, 4)$. The latter is referred to as a 'joint probability' because it consists of two component events.

Under the assumption that the two random variables involved in a joint event are independent, the probability of the joint event is obtained by multiplying the probability of the components. Thus

$$P(3, 4) = P(3) \times P(4)$$

In assuming that the dice are fair, we are stipulating that all of the $6 \times 6 = 36$ outcomes are equally likely. Since there are exactly six ways in which 7 can be rolled, it follows that $P(7) = 1/6$.

Thus far we have dealt with joint probabilities and their simple components, but many of the most important probabilities in natural language processing are conditional. A 'conditional probability' is one where attention is focused on a particular subset of the space of possible events. In the Conan Doyle story, as in English generally, it turns out that some words are extremely predictable from preceding context. The conditional probability focuses only on those cases where the first word is *Sherlock* and asks about the proportion of these cases where the next is *Holmes*. In fact, every time that the word *Sherlock* appears, the next word is *Holmes*. We notate this as:

$$P(\text{Holmes}_{n+1}|\text{Sherlock}_n) = 1$$

using Sherlock_n to denote the event that the nth word is *Sherlock* and Holmes_{n+1} to denote the event that the $n + 1$th word is *Holmes*. The vertical bar signals the conditioning of the probability and is read as 'given.' Notice in Table 1 that $P(\text{Sherlock}_n|\text{Holmes}_{n+1})$ is not the same thing. *Sherlock* occurs seven times only, always followed by *Holmes*, whereas *Holmes* occurs 46 times, as shown in the table. Therefore $P(\text{Sherlock}_n|\text{Holmes}_{n+1}) = 7/46$.

An important property of conditional probability is that it participates in the chain rule, which states that

$$P(A,B) = P(A) \times P(B|A) = P(B) \times P(A|B)$$

For the running example:

$$\begin{aligned} &P(\text{Holmes}_{n+1}, \text{Sherlock}_n) \\ &\quad = P(\text{Sherlock}_n) \times P(\text{Holmes}_{n+1}|\text{Sherlock}_n) \\ &\quad = P(\text{Holmes}_{n+1}) \times P(\text{Sherlock}_n|\text{Holmes}_{n+1}) \end{aligned}$$

Given that $P(\text{Holmes}) = 46/7051$ and $P(\text{Sherlock}) = 7/7051$, the reader will be able to verify that the two apparently different decompositions of the joint event lead to $P(\text{Sherlock, Holmes}) = 7/7051$.

Table 1 Occurrences of *Holmes* in 'A case of identity'

Combination		*No. of occurrences*
Mr	Holmes	14
Sherlock	Holmes	7
Windibank	Holmes	1
and	Holmes	1
at	Holmes	1
crushed	Holmes	1
remarked	Holmes	2
said	Holmes	18
while	Holmes	1

This equivalence is of more than academic interest in natural language processing, because it establishes a relationship between *P(A|B)* and *P(B|A)*.

$$P(B|A) = \frac{P(B) \times P(A|B)}{P(A)}$$

This relationship is known as 'Bayes's rule.' In natural language processing it is very common for *A* to be an 'observation' such as a word, whereas *B* is a 'hidden variable,' such as the part-of-speech tag for that word. The usefulness of the rule for natural language processing comes from the following facts:

- In many applications it suffices to find the value of *B* that maximizes $P(B|A)$, and the probability itself is just a means to this end. For purposes of maximization, we can simply drop $P(A)$ because it is a constant. When determining a part of speech, given a word, we do not care how probable the word itself is.
- It is often easier to collect the data and justify the assumptions that would be needed to form a sensible estimate of $P(B)$ and $P(A|B)$ than it would be for $P(B|A)$.

In beginning to talk of estimates, we are moving from probability to statistics, which is the topic of the next subsection.

Statistics

Imagine repeatedly tossing a coin and keeping track of heads and tails. It is reasonable to suppose that the successive trials are unaffected by the outcome of previous trials. It is also reasonable to expect that, over time, an unbiased coin will give approximately equal numbers of heads and tails. It is also true that, for a biased coin, the ratio between heads and tails will eventually settle down to some stable value that depends on how much material has been shaved off the edge of the coin by the person who biased it. In other words, given enough data we stand a good chance of obtaining a stable estimate of the true underlying value of the coin's bias. The same is true of language: if we see a large-enough training corpus, we will be in a position to estimate the underlying parameters of various different statistical models that we might assume to have generated that corpus.

For the Conan Doyle story, sensible probabilities can be obtained by counting the number of times that the relevant words occur in sequence. Although this process gives complete information about the specifics of 'A case of identity,' corpus researchers are typically interested in moving from the specific to more general hypotheses. For example, they may wish to make claims about the patterns of word occurrence in all Sherlock Holmes stories, in the whole of Conan Doyle, in Victorian and Edwardian literature, or in written English. Such claims rely not only on correct use of probability but also on scientific judgments about the extent to which small samples from some process can be regarded as representative of the process as a whole. The mathematics and science behind these judgments are at the heart of the field of statistics. Taken together, probability and statistics provide recipes for collecting data that one hopes will be representative, reducing that data to informative summaries and reasoning about the results. A good introduction to probability and statistics is DeGroot (1986).

Information Theory 'Information theory,' the mathematics of telegraphy, is central to statistical language modeling. It makes the idea of 'quantity of information' precise. This enables meaningful comparisons between different natural language processing systems.

Imagine watching a sequence of symbols go past on a ticker-tape. You have seen the symbols w_1 to w_{i-1} and you are waiting for w_i to arrive. You ask yourself the following question: *How much information will I gain when I see* w_i*?* Another way to express the same thing is: *How predictable is* w_i *from its context?* Each time we see a symbol we are more surprised or less surprised, depending on which symbol turns up. Large information gain goes with extreme surprise. If you can reliably predict the next symbol from context, you will not be surprised, and the information gain will be low. The average information gain is called 'entropy.' The entropy will be lowest when you can usually predict the next symbol with good reliability and highest when you often cannot. The amount of information gained when we see a w that has probability $P(w)$ turns out to be the logarithm of the probability (notated 'log $P(w)$'). If logarithms are taken to base 2, information is denominated in 'bits,' but if natural logarithms are used, the information is denominated in 'nats.'

A good statistical model of language is one that provides reliable predictions. It therefore tends to minimize entropy. The model that we have just described is a 'word-level language model.' Such a model would be useful for a speech recognition system that wishes to constrain its search for the next word by focusing attention on those that are likely to follow the words it has already recognized.

Cross-Entropy In the previous section we developed the idea that entropy is a measure of the average information gain from seeing the next symbol of a

ticker-tape. Now we imagine that we are still watching a ticker tape, whose behavior is still controlled by some probability distribution, but that we have imperfect knowledge of the true probabilities. Call the true probability model P and our imperfect knowledge of it P_M. When we see w we assess our information gain as log $P_M(w)$, not as the correct log $P(w)$. The average information gain that we assess in this situation is called the 'cross-entropy' of the signal. It is defined in terms of the imperfect model P_M as well as the true model P. It is a remarkable and important fact that the cross-entropy with respect to any incorrect probabilistic model is greater than the entropy with respect to the correct model.

This fact is important because it provides us with a justification for using cross-entropy as a tool for evaluating models. The search for a good model can be organized in the following way:

- Initialize the model with random (or nearly random) parameters.
- Measure the cross-entropy.
- Alter the model slightly (maybe improve it).
- Measure again, accepting the new model if the cross-entropy has improved.
- Repeatedly alter the model until it is good enough.

If we can guarantee that the alterations to the model will improve cross-entropy, so much the better, but even if some of the changes worsen the cross-entropy, the search may eventually yield good models.

Statistical Modeling

Linguistic phenomena are complex and involve much fine detail. In principle, given sufficient data, it would be possible to build accurate statistical models of such phenomena. Unfortunately, we never have sufficient data, a difficulty that is referred to as the 'sparse data problem.' It is counterintuitive but true that even millions of words of training data can be insufficient. Therefore we are driven to compromise, choosing to build relatively simple models that capture some but not all of the detail. The advantage of such models is that they can be trained on feasibly available data, the corresponding disadvantage being that they miss some of the potentially useful and interesting regularities that we would like to find.

Engineers and statisticians have a standard approach to this problem. The mathematics of this approach is straightforward, but its application to any particular modeling problem requires considerable insight and good judgment if it is to lead to effective solutions.

The easy part is the following: we begin by describing the situation that we wish to model as a collection of random variables. In the case of a word-level language model, we might note that the ith word w_i can often be predicted from its predecessors $w_1 \ldots w_{i-1}$. Using the chain rule, we obtain the following exact expression for the probability of a string as a product of simpler terms:

$$P(w_1 \ldots w_n) = P(w_1) \times P(w_2|w_1) \times P(w_3|w_1, w_2) \ldots \times P(w_n|w_1 \ldots w_{n-1})$$

or, using product notation

$$P(w_1 \ldots w_n) = P(w_1) \prod_{i=2}^{n} P(w_i|w_1 \ldots w_{i-1})$$

Unfortunately, while the first few terms of this series are simple, the later terms involve probabilities that are conditioned on large amounts of prior context. If a sentence is 20 words long, the prediction of the 20^{th} word will be conditioned on the preceding 19 words. It is very unlikely that this precise sequence of words will have occurred in the training data, so the model will have no basis for making a prediction. Even if the sequence has occurred, it is unlikely to have occurred often enough and with a broad enough range of different continuations for the predictions to be reliable. For example, if the sequence of 19 words has been seen exactly once in the training data, the model will have seen only one word after it and will predict that word with absolute certainty, excluding all others. Clearly the model has not seen enough data to warrant such certainty.

The general prescription for dealing with such situations is to adopt independence assumptions. This is done by ignoring some of the dependencies between the conditioning context and the variable whose value we are attempting to predict. Of course, we wish to keep the most important dependencies and drop the unimportant ones. For a word-level language model, it is plausible that the most important influence comes from the immediately preceding word, so we can adjust the model to neglect all dependencies that occur at a longer range than this. That is, we write:

$$P(w_1 \ldots w_n) \approx P(w_1) \times P(w_2|w_1) \times P(w_3|w_2) \ldots \times P(w_n|w_{n-1})$$

or, using product notation:

$$P(w_1 \ldots w_n) \approx P(w_1) \prod_{i=2}^{n} P(w_i|w_{i-1})$$

This greatly reduces the amount of data that must be stored. Where we previously needed to keep track of the successors of every contiguous subsequence of the training data, we can now get away with recording only the statistics about single words

and their possible successors. This alone makes the difference between feasibility and infeasibility. In addition, the chances of having enough training data to collect reliable statistics are greatly increased. The general strategy of simplifying probabilistic models by dropping the less important conditioning information is frequently very effective, but judgment is needed to decide how to identify that less important information and how to formulate a model that strikes the appropriate balance between feasibility and detailed predictive power. For the word-level language model, it turns out that the model outlined above (termed a 'bigram' model, because it needs only statistics over two-word sequences) is an effective compromise. A 'trigram' model (one that uses three-word sequences) is even better and is widely used in speech recognition.

As we saw in the previous paragraph, restrictive independence assumptions can reduce the impact of the sparse data problem by reducing the number of bins into which the limited amount of training data needs to be sorted. However, it can still happen that the finiteness of the training data cause problems. Specifically, in the case of the word-level language model, when we attempt to use it to predict new sentences, it may encounter a bigram that it has never seen. Unadjusted, the model would give any sentence containing such a bigram a probability of zero, regardless of what else is going on in the sentence. This is usually not what is required: one wants to be able to make some kind of discrimination between such sentences, so realistic models involve 'smoothing.' The idea is to slightly **discount** the evidence from the training corpus, so that unexpected events can be given some minimal probability. One simple way to do this is to form a smoothed model by averaging together the bigram language model with an extremely noncommittal model, such as the one that says that all words are equally likely to occur. This latter is the 'uniform' model. Since one does trust the bigram language model to some extent, and one wants the result to be informative, one will give the bigram model greater weight in the averaging process than the uniform model. Although sentences with zero-frequency bigrams will still get low probabilities, these will not be zero. Other, more complex, smoothing schemes are described in chapter 6 of Manning and Schütze (1999). Another approach to the sparse data problem is to use a model of language based on parts of speech, as suggested in the earlier discussion of Chomsky. Once again, an option is to average together the more robust parts-of-speech language model with the more delicate word-based language model to yield a combined model that is both reasonably robust and reasonably delicate.

Machine Learning for Natural Language Processing

Given the availability of large-scale language data, it becomes possible to apply standard algorithms from artificial intelligence and data mining. The most relevant sub-field of artificial intelligence is called 'machine learning.' Since this activity is now so widespread in natural language processing, it is useful to review a few standard concepts from the machine learning literature that are taken as read in the application of machine learning to natural language processing.

Machine learning algorithms come in two main flavors. The first type is 'supervised.' These algorithms learn from pairs of a predictive context x and a known correct answer y. For example, the algorithm designed to disambiguate the senses of words might be shown a large number of contexts in which the word *line* occurs with the different senses listed below, which are selected from the WordNet electronic database (Fellbaum, 1998).

1. *line* – (a mark that is long relative to its width)
2. *line* – (text consisting of a row of words written across a page or computer)
3. *line* – (a conductor for transmitting electrical or optical signals or electric power)

It is likely that the dominant sense in geometry textbooks is the first, that the second sense is frequent in editorial instructions to authors, and that the third is well-represented in telephone repair manuals. In any event, the supervised learning paradigm assumes that some expert has marked up instances of *line* with the appropriate sense. The business of the supervised learner is to learn the relationship between visible clues in the surrounding text and the correct answers. Methods for doing this are referred to as 'classifiers.' Many classifiers are used in natural language processing, including the 'Naive Bayes classifier,' which often performs well in spite of making extremely bold independence assumptions, and 'decision-tree classifiers,' which work by designing a structured series of questions to ask about clues that might or might not be present in the predictive context. For example, a decision tree trying to disambiguate *line* might begin by asking whether the word *angle* appears in the predictive context. If the answer is *yes,* it might conclude that we are probably dealing with the first sense. But if not, it might proceed to ask whether the word *transmission* is present, and so on. Algorithms exist for constructing decision trees directly from annotated corpora, without recourse to human intervention.

The second type of machine-learning algorithm used in natural language processing is the 'unsupervised' algorithm. In this setting we do not assume that correct answers are available, and instead try to solve the problem of grouping the predictive contexts x into coherent groups. Methods for doing this are usually referred to as 'clustering algorithms.' An effective clustering algorithm might group together all the samples collected from geometry textbooks, but further processing would be needed to recognize that the instances of *line* in these samples are predominantly of the first of the WordNet senses listed. Such algorithms are useful when working with unannotated corpora, but the groups that they form are sometimes difficult to interpret and criteria for success are not as obvious as in the case of supervised algorithms. Manning and Schütze (1999) cover both supervised and unsupervised word-sense disambiguation in chapter 8 of their textbook.

Modeling of Sequences

Our discussion of word-level language models has touched on one of the key distinguishing features of natural language processing as an application area for machine learning. The crucial aspect of these models was that the desired output was not a single word but a sequence of interrelated words. This section describes statistical approaches to one problem that has this character: the assignment of parts of speech. The discussion is necessarily brief, but is intended to indicate the nature of the particular challenge that is posed by the task of predicting structured outputs.

The structured prediction task we consider is 'part-of-speech tagging' (*see* **Part-of-Speech Tagging**). In this task the goal is to assign part-of- speech labels to words of a running text. The basis for one solution to this problem is the 'hidden Markov model' (HMM). This is an enrichment of the word-based language model we saw earlier. In this model, the sequence of observable words is thought of as associated with a parallel sequence of unobserved parts of speech. Our task, since we see only the words, is to recover a best guess at the likely sequence of parts of speech. Hidden Markov models can be dramatized via a well-known 'urns and balls' metaphor. In the urns and balls model, we assume that there are N large urns in a room. Each urn contains a number of colored balls. There are M different colors, and the relative proportion of each color may differ from urn to urn. A genie randomly jumps from urn to urn, randomly choosing a ball from each of the urns that it visits. It shouts out the color of the ball, then replaces it in the urn. The observer hears the shouts but cannot see where the genie is and does not know which urn the genie is currently in. If one knows the contents of the urns, there are good algorithms that allow a listener aware only of the genie's shouts to form a rational best guess about the sequence of urns that the genie is most likely to have visited.

This metaphor can be adapted to the part-of-speech assignment problem. Each urn is identified with a single part of speech. There may be one urn for nouns, another for verbs, another for prepositions. Instead of colored balls, the urns contain words. Among the contents of the urn for verbs will be words like *devour, react,* and *solve,* which are unambiguously verbs. They will be present in the urn in proportion to the frequency with which they occur in the language. Similarly, unambiguous nouns such as *door* and *penguin* will be present only in the noun urn. A listener who hears any of the aforementioned words can be sure from which urn the genie drew that word.

But words with part-of-speech ambiguities afford no such certainty. If the genie shouts out *line,* which is ambiguous between noun and verb, the best that the listener can do is to take account of the conditional probabilities $P(\text{line}|\text{noun})$ and $P(\text{line}|\text{verb})$ to form a best guess. These probabilities are called 'emission' probabilities and are relevant to the competition between noun and verb as ways for explaining the presence of the word *line*.

With luck, the listener may also have noticed that the previous word was *the* and that the urn for articles is very far from the urn for verbs but quite close to the urn for nouns, making it unlikely that the genie will have just jumped all the way from article to verb, but relatively likely that it has jumped from article to noun. This fact about the layout of the room could be encoded in the form of 'transition probabilities' – $P(\text{noun}|\text{article})$ and $P(\text{verb}|\text{article})$ – which tell us which part of speech is likely to follow which. The algorithms that allow a best guess about the sequence of jumps taken by the genie also produce a coherent best guess at the likely sequence of parts of speech that goes with a sequence of words. Remarkably, it also turns out that statistical models for the part-of-speech assignment problem can be learned in unsupervised fashion from unannotated text by using a variant of the scheme for reducing cross-entropy that was outlined earlier.

More details of the training and use of HMMs are given by Huang *et al.* (1990). This monograph focuses mainly on speech recognition, but the same algorithms apply to the part-of-speech assignment problem. The treatment of HMMs (sequential structure) has a parallel for probabilistic context-free grammars (tree structure). This cannot be covered

here but is dealt with in detail in chapter 12 of Jurafsky and Martin (2000). The same theoretical possibilities in terms of training from data and recovery of a best guess are available even for these more complex structures.

It is tempting to suppose that Swift's literary engine was a primitive part-of-speech tagger, that the 40 iron handles controlled the choice of appropriate urns, that the "Tenses, Moods, and Declensions" were encoded into the parts of speech corresponding to the urns, and that the chance of producing semi-coherent English text from it was greater than Gulliver was able to appreciate. Literary merit is another matter.

Conclusion

Statistical processing of language is a well-established subfield of language science, drawing on linguistics, computer science, and engineering. For extended introductions to the field, see Manning and Schütze (1999) and Jurafsky and Martin (2000), two textbooks that do a good job of representing the breadth of the field. The increasing availability of language samples in electronic form opens many possibilities for both language technology and theoretical investigation. These samples include carefully curated collections that have been annotated by expert linguists, larger but more lightly curated collections that have been annotated by semiautomatic means, and huge bodies of uncurated text culled from the World Wide Web. It is clear that all these are potentially useful sources of linguistic knowledge, but it is less clear how these sources might be used for maximum scientific and technological benefit.

See also: Classification of Text, Automatic; Corpora; Corpus Linguistics; Natural Language Processing: Overview; Parsing: Statistical Methods; Part-of-Speech Tagging; Part-of-Speech Tagging; Text Mining; Treebanks and Tagsets; WordNet(s).

References

Abney S (1996). 'Statistical methods and linguistics.' In Klavans J & Resnik P (eds.) *The balancing act: combining symbolic and statistical approaches to language.* Cambridge, MA: The MIT Press. 1–26.

Bod R, Hay J & Jannedy S (2004). *Probabilistic linguistics.* Cambridge, MA: MIT Press.

Chomsky N (1957). *Syntactic structures.* The Hague: Mouton.

DeGroot M (1986). *Probability and statistics* (2nd edn.). Reading, MA: Addison-Wesley.

Fellbaum C (ed.) (1998). *WordNet: an electronic lexical database.* Cambridge, MA: MIT Press.

Huang X, Ariki Y & Jack M A (1990). *Hidden Markov models for speech recognition.* Edinburgh: Edinburgh University Press.

Jurafsky D & Martin J (2000). *Speech and language processing.* Upper Saddle River, NJ: Prentice-Hall.

Kučera H & Francis W (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Manning C & Schütze H (1999). *Foundations of statistical natural language processing.* Cambridge, MA: MIT Press, 1999.

Mosteller F & Wallace D (1964/1984). Applied Bayesian and classical inference: the case of *The Federalist Papers* (2nd edn.). New York: Springer.

Pereira F (2000). 'Formal grammar and information theory: together again?' *Philosophical Transactions of the Royal Society 358,* 1239–1253.

Language Socialization

P B Garrett, Temple University, Philadelphia, PA, USA

Introduction

Language socialization is the process whereby a child or other novice develops communicative competence through interactions with older and/or more experienced persons. Language socialization researchers seek to understand how novices learn to use language in the culturally specific ways that enable them to participate competently in the social life of a particular community. Another, equally important goal of language socialization research is to understand the broader sociocultural contexts within which this developmental process occurs, from local to global levels of analysis.

Origins

The language socialization research paradigm was initially formulated in the early 1980s by Elinor Ochs and Bambi Schieffelin (Ochs and Schieffelin, 1984; Schieffelin and Ochs, 1986a, 1986b). At that time, a significant body of research on language acquisition already existed, as did another on socialization

here but is dealt with in detail in chapter 12 of Jurafsky and Martin (2000). The same theoretical possibilities in terms of training from data and recovery of a best guess are available even for these more complex structures.

It is tempting to suppose that Swift's literary engine was a primitive part-of-speech tagger, that the 40 iron handles controlled the choice of appropriate urns, that the "Tenses, Moods, and Declensions" were encoded into the parts of speech corresponding to the urns, and that the chance of producing semi-coherent English text from it was greater than Gulliver was able to appreciate. Literary merit is another matter.

Conclusion

Statistical processing of language is a well-established subfield of language science, drawing on linguistics, computer science, and engineering. For extended introductions to the field, see Manning and Schütze (1999) and Jurafsky and Martin (2000), two textbooks that do a good job of representing the breadth of the field. The increasing availability of language samples in electronic form opens many possibilities for both language technology and theoretical investigation. These samples include carefully curated collections that have been annotated by expert linguists, larger but more lightly curated collections that have been annotated by semiautomatic means, and huge bodies of uncurated text culled from the World Wide Web. It is clear that all these are potentially useful sources of linguistic knowledge, but it is less clear how these sources might be used for maximum scientific and technological benefit.

See also: Classification of Text, Automatic; Corpora; Corpus Linguistics; Natural Language Processing: Overview; Parsing: Statistical Methods; Part-of-Speech Tagging; Part-of-Speech Tagging; Text Mining; Treebanks and Tagsets; WordNet(s).

References

Abney S (1996). 'Statistical methods and linguistics.' In Klavans J & Resnik P (eds.) *The balancing act: combining symbolic and statistical approaches to language.* Cambridge, MA: The MIT Press. 1–26.
Bod R, Hay J & Jannedy S (2004). *Probabilistic linguistics.* Cambridge, MA: MIT Press.
Chomsky N (1957). *Syntactic structures.* The Hague: Mouton.
DeGroot M (1986). *Probability and statistics* (2nd edn.). Reading, MA: Addison-Wesley.
Fellbaum C (ed.) (1998). *WordNet: an electronic lexical database.* Cambridge, MA: MIT Press.
Huang X, Ariki Y & Jack M A (1990). *Hidden Markov models for speech recognition.* Edinburgh: Edinburgh University Press.
Jurafsky D & Martin J (2000). *Speech and language processing.* Upper Saddle River, NJ: Prentice-Hall.
Kučera H & Francis W (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.
Manning C & Schütze H (1999). *Foundations of statistical natural language processing.* Cambridge, MA: MIT Press, 1999.
Mosteller F & Wallace D (1964/1984). Applied Bayesian and classical inference: the case of *The Federalist Papers* (2nd edn.). New York: Springer.
Pereira F (2000). 'Formal grammar and information theory: together again?' *Philosophical Transactions of the Royal Society 358,* 1239–1253.

Language Socialization

P B Garrett, Temple University, Philadelphia, PA, USA

Introduction

Language socialization is the process whereby a child or other novice develops communicative competence through interactions with older and/or more experienced persons. Language socialization researchers seek to understand how novices learn to use language in the culturally specific ways that enable them to participate competently in the social life of a particular community. Another, equally important goal of language socialization research is to understand the broader sociocultural contexts within which this developmental process occurs, from local to global levels of analysis.

Origins

The language socialization research paradigm was initially formulated in the early 1980s by Elinor Ochs and Bambi Schieffelin (Ochs and Schieffelin, 1984; Schieffelin and Ochs, 1986a, 1986b). At that time, a significant body of research on language acquisition already existed, as did another on socialization

(sometimes referred to as enculturation). But the two had developed quite separately from one another. Language acquisition research, rooted in developmental psychology and psycholinguistics, tended to treat language acquisition as a rather self-contained individual developmental process, largely ignoring the sociocultural contexts within which it occurs; conclusions drawn from studies conducted in mainstream North American and European settings were assumed to be universally valid. Meanwhile socialization research, rooted in anthropology and sociology, was conducted in a variety of ethnographic settings worldwide, but gave little attention to the central role of language as the primary medium through which socialization occurs. Working in collaboration with researchers from several disciplinary backgrounds (including linguistic anthropology, developmental psychology, sociolinguistics, applied linguistics, and education), Schieffelin and Ochs sought to combine the strengths of both of these established bodies of research and to bridge the gap between them.

Axioms and Aims

As this synthetic approach suggests, a central assumption in language socialization research is that the acquisition of language is inseparable from the acquisition of other kinds of cultural knowledge and practices. This is perhaps nowhere more clearly seen than in interactions between young children and their caregivers, which were the main focus of the earliest language socialization studies (Schieffelin and Ochs, 1986a, 1986b). As a young child is acquiring language, she or he is simultaneously developing a repertoire of social skills and a culturally specific world view. In learning how to use the language(s) of their community, children also learn how to think, how to comport themselves, even how to feel in particular situations and how to express (or otherwise manage) those feelings. As a developmental process, then, language acquisition is far more than a matter of learning to produce grammatically well-formed utterances. It is also a matter of learning how to use language in socially and pragmatically appropriate, locally intelligible ways, and as a means of engaging with others in culturally meaningful activities. Over the course of time, this makes it possible for the child to engage with others in an increasingly broad range of social contexts, and to assume increasingly complex roles and identities. (To varying degrees, these assumptions are shared by researchers in allied fields such as developmental psychology and developmental pragmatics; see Bloom [1998] and Bugental and Goodnow [1998] for useful overviews, and see Blum-Kulka and Snow [2002] for a recent collection of case studies.)

Language socialization research is thus concerned with the microgenesis of communicative competence (Schieffelin and Ochs, 1996), which comprises but also goes well beyond linguistic competence in the generativist sense. Communicative competence also comprises the practical knowledge, much of it prediscursive (if not preconscious), that one must have in order to use language as a social tool, to engage in talk as a social activity, and to coconstruct meaningful interactive contexts with others. Language socialization research is thus concerned with all of the knowledge, orientations, and practices that make it possible for an individual to function as – and, crucially, to be regarded by others as – a competent member of (or participant in) a particular community, however broadly or narrowly defined.

Examples of Key Insights and Areas of Investigation

Since the initial formulation of the language socialization research paradigm in the early 1980s, language socialization studies have been conducted in a broad range of settings worldwide. These studies have provided the empirical foundations for various kinds of comparative insights, which in turn have given rise to new theoretical developments and productive new areas of thematic focus as the paradigm has continued to develop.

Nonuniversality of Baby Talk

In their pioneering comparison of language socialization practices among members of three groups – white middle-class Americans, Kaluli (of Papua New Guinea), and Samoans – Ochs and Schieffelin (1984) found important culturally based differences in several areas. One such difference concerns baby talk – the simplified, stylized register that caregivers in some groups (e.g., white middle-class Americans) use with infants and young children. Previously, language acquisition researchers had assumed, based on studies conducted in mainstream North American and European contexts, that baby talk is universal (i.e., used by caregivers in all societies), and that this ostensibly accommodative way of speaking facilitates the child's acquisition of language. Ochs and Schieffelin found that not only do Samoans and Kaluli never use baby talk; they rarely address preverbal children at all, and do not treat them as conversational partners. Even so, young Kaluli and Samoan children are immersed in a richly verbal environment from the earliest days of their lives. Ochs and Schieffelin found that they

acquire language at more or less the same age, and at the same developmental rate, as children in Western societies.

Differing Cultural Understandings of Human Development

Ochs and Schieffelin also found that the preverbal child's abilities and tendencies are conceptualized quite differently from one group to another, and that the professed goals of language socialization practices differ accordingly. The egalitarian Kaluli regard preverbal children as 'soft' and helpless, based on not just their lack of physical abilities, but also their inability to use language as a social tool, i.e., as a means of getting what they need or want from others (Schieffelin, 1990). For the Kaluli, a primary goal of socialization is to 'harden' children, which is largely a matter of helping them learn to use 'hard' (adult-like, socially effective) language. (From a Kaluli perspective, the use of baby talk would be antithetical to this goal.) Samoans, in contrast, regard preverbal infants as willful, defiant, and antisocial; among their foremost concerns is that children learn their place in hierarchical Samoan society, and that they learn to act and speak accordingly (Ochs, 1988). Age is an important dimension of social stratification in Samoa, so in this regard young children are near the bottom of the hierarchy. They must learn to adapt themselves to social situations, as they cannot expect their elders to accommodate them (such as by using deliberately simplified language). The first words attributed to children in these two societies reflect these differing cultural conceptions of the nature of the child: Kaluli believe that a child's first meaningful utterances will be the words for *mother* and *breast*, while Samoans expect to hear a defiant *(Eat) shit!* (Ochs and Schieffelin, 1984).

Influence of Local Social Norms on Children's Acquisition and Production of Specific Linguistic Forms

Language socialization researchers have demonstrated that specific aspects of culture and social organization such as those just described can influence children's acquisition of language in quite specific ways at the level of linguistic form. A striking example is Platt's (1986) study of Samoan children's acquisition of the deictic verbs *sau* 'to come' and *'aumai* 'to bring/give.' Young children frequently hear both verbs used as imperatives in everyday contexts, and *sau* is the less semantically complex of the two. Based on this information alone, the expectation would be that children would begin using *sau* productively earlier than they would begin using *'aumai*. But as noted above, young Samoan children occupy a low position in an age-stratified social hierarchy. Very rarely is it appropriate for them to tell anyone to come – it is the prerogative of higher-status persons in Samoan society to remain stationary and to have lower-ranking persons come to them. Young children are expected and encouraged to appeal to higher-status persons to give them things (such as food), however. Thus the number of persons toward whom a young child can appropriately use the imperative *give* is far greater than the number of persons whom the child can tell or ask to *come*. Young children's much more frequent use of *'aumai* 'give' can be taken as evidence that they learn the social norms that organize and constrain the use of these two verbs at the same time as they learn the linguistic forms.

Differing Social Functions of Specific Verbal Practices across Cultures and Social Groups

Language socialization researchers have found that the same basic type of practice or activity may serve different functions in different communities, even in the same society. A good example is verbal teasing. Teasing may be used as a means of social control, i.e., to shame a child and thereby discourage him or her from engaging in a particular behavior (perhaps as an alternative to corporal punishment or some other form of coercive physical intervention, which may be dispreferred); to toughen the child by teaching him or her how to deal with others' affronts, and how to be self-assertive in return; or as a form of interactive play that provides a basis for a child and an adult to engage each other verbally without need or expectation of exchanging substantive information (Schieffelin and Ochs, 1986b). Examples such as this reveal the potential of language socialization research to reveal both universal and culturally specific aspects of socialization practices, and of communicative practices more generally.

Delimiting Universals of Communicative Practice

Ways of dealing with unintelligibility provide another such example and may offer insight into a universal aspect of communicative practice. Schieffelin and Ochs (1996) propose a universal set of possible responses to unintelligibility, first making a distinction between speaker-rooted unintelligibility (in which the speaker perceives his or her own utterance to be unintelligible to an interlocutor) and addressee-rooted unintelligibility (in which the speaker perceives an interlocutor's utterance to be unintelligible). In each case, they assert, four basic responses are available to speakers everywhere. But in different communities, one response may be preferred over the other three; or two or three may be regarded as more or less equally valid options while another is strongly dispreferred.

In the case of addressee-rooted unintelligibility, the four universal options are:

1. ignore unintelligibility;
2. display nonunderstanding;
3. verbally guess at what interlocutor might be saying;
4. negatively sanction interlocutor's unintelligibility (e.g., by teasing or shaming).

Schieffelin and Ochs observe that among Kaluli and Samoans, option #3 is least preferred – quite unlike in mainstream American society. Looking beyond these three groups, one might reasonably surmise that in East Asian societies such as Japan and Korea, where the social dynamics of honorification are pervasive and subtle, empathetic and intuitive styles of interaction are strongly preferred, and there is much concern with face (one's own as well as that of one's interlocutor), option #4 would be dispreferred in most situations; whereas #1, #2, and #3 would all be more preferred, but to differing degrees under different social circumstances (e.g., #3 is much more likely to be used in casual conversation with a peer than in a workplace exchange with one's supervisor).

Language Practices at Home and in School

Certain themes addressed in early language socialization research have been carried forward and taken in new directions in subsequent years. Heath's (1983) comparative study of language use in two working-class communities in the USA (one black, one white) and the implications for children's rates of success in school has had enduring influence. Heath's work continues to inform recent studies that examine various kinds of continuities and disjunctures between children's home environments and the classroom. The dinner table conversations of white middle-class American families (Ochs *et al.*, 1989, 1992), for example, have been shown to foster specific types of problem-solving orientations and to encourage children to display skills that are expected and rewarded in the classroom. In other settings, cultural disjunctures between home and school lead to poor educational outcomes (Watson-Gegeo, 1992). Research on children from immigrant and other nonmainstream communities in the USA suggests that when classroom activities and modes of interaction draw on communicative practices and participant structures that are familiar to these children, their levels of participation and academic achievement improve significantly (Gutiérrez *et al.*, 2001). Heath's influence can also be traced to other recent studies that focus on the role of narrative in language socialization. Baquedano-López (2001), for example, examines how narrative practices serve as a resource for the socialization of a transnational Mexican identity in a Catholic parish in Los Angeles; and Capps and Ochs (1995) consider the central role that narratives of personal experience play in the discursive construction, and quite possibly the social reproduction, of agoraphobia.

The Development of Subjectivities

A central theme of virtually all language socialization studies is the development of locally intelligible subjectivities, or ways of being in the social world (Kulick and Schieffelin, 2004). This is clearly seen, for example, in Clancy's (1999) study of how affective states and verbal expressions of affect are negotiated between Japanese mothers and their young children; in Schieffelin's (1990) examination of how Kaluli mothers cultivate sibling relationships among their children such that elder sisters will 'feel sorry for,' and always be willing to give to, their younger brothers; and in Fader's (2001) study of how Hasidic Jews in New York City use literacy practices in socializing girls' gender and ethnic identities – also an important means by which symbolic boundaries separating the Hasidim from other groups (including other Jewish groups) are maintained. As these examples suggest, virtually all aspects of subjectivity, including affect, morality, and desires, are shaped in culturally specific ways – and in accordance with cultural preferences, by and large – through language socialization. But in an interesting and productive reorientation of this perspective, Kulick and Schieffelin (2004) urge language socialization researchers to consider 'bad subjects' – those individuals in every community who persistently display culturally dispreferred traits and/or engage in nonnormative, deviant behaviors. As Kulick and Schieffelin point out, "the focus on expected and predictable outcomes is a weakness if there is not also an examination of cases in which socialization doesn't occur, or where it occurs in ways that are not expected or desired" (Kulick and Schieffelin, 2004: 355). Language socialization research must account for reproduction as well as "why socializing messages to behave and feel in particular ways may also produce their own inversion" (Kulick and Schieffelin, 2004: 356).

Languages and Cultures in Contact

As noted previously, recent language socialization studies explore the ways in which locally situated interactions articulate with macro-level social processes, such as cultural revitalization movements and postcolonial nation-building. Many of these studies have been conducted in settings characterized by sustained contact between languages and cultures, such as occurs in postcolonial, creole, diasporic, urban, and border communities (Baquedano-López, 2001; Fader,

2001; Garrett, 2005; Kulick, 1992; Moore, 2004; Paugh, 2005). These studies investigate the ways in which language socialization processes unfold in linguistically and socioculturally heterogeneous settings characterized by bilingualism and multilingualism, code-switching, language shift, and other contact-induced linguistic and sociocultural phenomena. The coexistence of two or more codes within a particular community, whatever the sociohistorical circumstances that have brought them into contact, is rarely a neutral or unproblematic state of affairs. It tends to be a focal point of discursive elaboration and social conflict, with complex linkages to other, equally contested issues that play out on multiple levels, from the household to the state. As they are socialized to use language, children are also socialized into knowledge of these intimately related issues, and of preferred and dispreferred ways of dealing with them, from their earliest years.

Dynamics of Language Shift

In some contact settings, language socialization practices may be an important mechanism of language shift – a point made vividly by Kulick's (1992) study of rapid shift in a small village in Papua New Guinea. Building on Kulick's work – which shows that local cultural and ideological factors may be of greater importance in accounting for language shift than the macrosociological (e.g., political and economic) factors that are more commonly invoked – more recent studies have shown language shift to be a contingent, nonlinear phenomenon in which language socialization practices play a crucial but often subtle mediating role. Paugh (in press) shows that young children in the Caribbean island of Dominica are encouraged by their parents and other adults to acquire English, which is contributing to language shift away from the local Afro-French creole. But in later years, as children spend increasing amounts of time interacting with peers beyond the supervision (and earshot) of adults, they increasingly use the creole in their play as a way of enacting adult roles and activities. Although ultimately it may not prevent language shift from running its course, at present such pretend play seems to provide older children with opportunities to develop some degree of proficiency in the creole. In his investigation of a quite similar situation in the Caribbean island of St. Lucia, Garrett (2005) examines code-specific genres as the basis for socialization activities in which adults encourage young children to be verbally self-assertive. Cursing, insulting, and other locally valued self-assertive ways of speaking conventionally require use of the historically stigmatized Afro-French creole language that many St. Lucian caregivers (much like their counterparts in Dominica) otherwise discourage children from using. The persistence of socialization routines in which adults playfully urge children to use the creole in self-assertive ways favors the maintenance of such code-specific genres, which in turn may be having a retarding or dampening effect on language shift.

The Persistence of Everyday Practices

These and other recent studies suggest that a language socialization approach can yield a more nuanced account of ongoing changes and shifts in local communicative practice than can larger-scale, quantitatively oriented studies that are less attentive to situated interactions (particularly those occurring at the critical juncture between generations). In a case where language shift is already quite far along, for example, a language socialization approach may reveal that culturally specific ways of using language persist in speakers' use of the 'new' code. In her study of a Navajo community, Field (2001) demonstrates that although today's bilingual caregivers often speak English to children, they continue to socialize traditional Navajo values of autonomy, self-determinacy, and respect through the use of a triadic participant structure in issuing directives. Based on this observation, Field (2001: 249–250) proposes that "certain aspects of language use may be more conservative, or more resistant to change, than code." She goes on to assert, "[I]t is exactly those aspects of a speech community's interaction that are tacitly taken for granted that are also the most basic, pervasive, and resistant to change. Furthermore, they are maintained through the most mundane routines and forms of everyday communicative practice – which also happen to be the preferred context for research on language socialization."

Essential Features of Language Socialization Research

In whatever setting a particular study is conducted, and whatever specific linguistic and sociocultural phenomena are the focus of the investigation, four key features are essential to language socialization research. These four features reflect the paradigm's interdisciplinary origins as well as its commitment to taking a maximally holistic perspective on the relationships among language, culture, and society.

1. **Longitudinal study design.** Language socialization researchers closely track developmental changes in individual subjects by periodically recording their participation in naturally occurring socialization interactions and activities over a developmentally significant span of time. In order for such tracking to be feasible, a language socialization study

usually focuses on a relatively small number of children or novices – typically three to six, or a single small cohort (if the study is conducted in a school or other institutional setting). Qualitative depth of analysis is emphasized over quantitative breadth. Data in the form of naturalistic audio and/or audio-video recordings of the focal children or novices interacting with caregivers and other community members are collected at regular intervals (e.g., monthly), usually over the course of a year or more of sustained fieldwork. Some researchers revisit their field sites periodically over several years or even decades (e.g., Schieffelin, 1996) in order to keep up with individual and community development over longer spans of time.

2. **Field-based collection and analysis of a substantial corpus of naturalistic audio or audio-video data.** Regular, periodic data collection as described above gives rise to a large corpus of recordings; a year-long study typically yields 75–100 hours. A corpus of this size strikes a balance between ethnographic and longitudinal adequacy and practical manageability. But collection of the recordings is only a first step; in order for them to serve as a meaningful data set, the researcher must transcribe and annotate them while in the field. This is accomplished with the aid of local consultants, usually members of the community in which the research is being conducted (such as older relatives of the focal children). For language socialization researchers, this one-on-one collaboration with local consultants is indispensable. In addition to assisting with the most basic aspects of transcription (such as clarifying specific words and phrases captured in the recordings), consultants can bring to the researcher's attention layers of meaning that would otherwise escape his or her notice or understanding. Collaborative transcription also provides ongoing opportunities for the researcher to benefit from consultants' native-speaker intuitions about the use of particular linguistic forms and variants, and their perspectives on many other aspects of local social life.
3. **A holistic, theoretically informed ethnographic perspective.** This is achieved in part through sustained fieldwork and a commitment to ethnographic methods (including participant observation), and in part through familiarity with current theoretical issues and debates concerning such methods. Both depth and breadth of ethnographic observation are important in language socialization research. In addition to tracking individuals over the course of time (as described above), the researcher must observe and record in a broad variety of contexts in order to understand how different social settings may influence those individuals' language usage and modes of participation. Doing so allows the researcher to observe and record a broad range of persons as well; in effect, tracking a particular focal subject across contexts provides access to an entire social network, and often to a broad cross-section of the community that includes fictive kin, peers, neighbors, etc. Although most recorded data are collected during everyday activities, the researcher must be attentive as well to exceptional events (i.e., those that occur rarely or unpredictably) and to periodic activities such as those associated with agricultural, political, and ritual cycles. The systematic collection of recorded data that is central to any language socialization study may be supplemented by surveys, interviews, elicitation sessions, or other methods, depending upon the kinds of data that are needed in order to address the study's central research questions. Whatever complementary methods are chosen, the researcher should have a thorough understanding of the theoretical issues in which they are based.
4. **Attention to both micro and macro levels of analysis, and to linkages between them.** This can be considered part of the ethnographic perspective outlined above, but is important enough to merit consideration in its own right. Language socialization research is not just a matter of producing detailed ethnographic accounts of individual developmental processes and the local contexts in which they occur. An overarching goal is to understand how such individual developmental processes relate to larger sociocultural and historical processes. As they analyze their recordings and other microethnographic data, language socialization researchers are constantly on the lookout for patterns and principles that may also be discernible at macro levels of analysis. Likewise, when they make macroethnographic observations, they consider various ways in which the patterns or principles identified may be writ small in their recorded data. Ultimately, the language socialization paradigm is comparative in perspective, recognizing that while some aspects of language socialization are universal, others vary considerably from one sociocultural setting (or historical period) to another. Attention to micro–macro connections is an important means by which researchers are able to distinguish between the universal and the culturally specific, and to consider the relationships between them.

In recent years, some researchers who have claimed to be doing language socialization research have

largely if not completely ignored one or more of these four essential areas. (See Kulick and Schieffelin [2004] and Watson-Gegeo and Nielsen [2003] for recent comments on language socialization research design and methodology.) A short-term study that offers no longitudinal observations, a study based solely on questionnaires and interviews, or one that makes use of naturalistic recordings but is not ethnographic in any meaningful sense simply does not constitute language socialization research. (Such a study may yield interesting data and may contribute meaningfully to other research paradigms, of course.)

That said, the language socialization paradigm is sufficiently flexible to comprise a broad range of studies that place varying degrees of emphasis on the four features outlined above. Since the paradigm's initial formulation, certain trends have emerged in language socialization research that reflect broader trends in the social sciences. Most early language socialization studies were conducted in small-scale non-Western communities and focused primarily on microethnographic levels of analysis. These studies yielded classic ethnographic descriptions of these communities as well as fine-grained, strongly longitudinal accounts of how individual developmental processes unfold within them. More recent studies, many of them conducted in Western societies and in various kinds of socioculturally and sociolinguistically heterogeneous settings, have tended to emphasize the ways in which individuals and local communities are implicated in macro-level processes such as those associated with globalization phenomena, and the ways in which everyday practices of individuals shape and are shaped by those processes (Garrett and Baquedano-López, 2002). Similarly, language socialization studies conducted by researchers of different disciplinary backgrounds show certain broad differences in orientation, such as whether they place greater emphasis on ethnographic topics and issues of social theory, or on grammatical development and the functions and distribution of specific linguistic forms and structures (Ochs and Schieffelin, 1995).

Language Socialization across the Lifespan

Language socialization is always a reciprocal, dialectical process in which the learner, regardless of age or level of experience, is much more than a passive recipient of input. The child or novice plays an active role in coconstructing every interaction in which she or he participates, however limited that participation might be; even a preverbal infant must be regarded as an emergent participant (de León, 1998), who, while being socialized himself or herself, is in various ways socializing others into such roles as mother, father, and elder sibling. Novices can often be observed to resist socialization, or to steer socializing interactions in new, sometimes unexpected directions. Older children may even assume the role of expert *vis-à-vis* their elders, as when a school-age child introduces his or her parents to a new technology (e.g., computers), or when adult immigrants with limited proficiency in the language of the host society rely on their bilingual children to assist them in dealing with persons and institutions outside the household.

Although the majority of studies conducted thus far have focused on young children, others, particularly in recent years, have focused on language socialization later in the life cycle: in middle childhood, adolescence, and adulthood. As this suggests, language socialization researchers consider socialization to be a lifelong process. Even those who study young children avoid treating child development as unilinear progression toward a static, monolithic adult status. Likewise, in studies of older children and adolescents, it is not assumed that they are gradually taking on pre-existing adult roles and identities; on the contrary, they can often be observed, particularly within their peer groups, to contest and renegotiate the tropes and discourses of identity (including gender and racial/ethnic stereotypes) that circulate among adults in their communities and in society at large.

Entry into adult status, however locally defined, is by no means the end of socialization. Adults continue to be socialized into new roles, statuses, identities, and practices, many of which involve new ways of using language. Adults may find it necessary or desirable to master new registers or styles associated with changes in their vocational or professional lives, or with new avocations or other activities that broaden their social horizons and involve participation in new communities of practice. Similarly, emigration, religious conversion, and other significant life changes may make it necessary or desirable for adults to master new codes and/or new discursive genres, which may involve either spoken or written forms of language (Schieffelin, 1996).

Reproduction and Continuity

For language socialization researchers, close analyses of individuals' participation in naturally occurring interactions are not ends in themselves, but provide empirical points of entry into larger issues of sociocultural reproduction and transformation. Among the most significant contributions of language socialization research is the insight that it yields into the everyday life of a community – the mundane activities and interactions in which ordinary individuals

participate on a daily basis, constituting the warp and weft of social life. As linguistic anthropologists have long recognized, a community's norms, values, ideologies, patterns of social organization, and cultural preferences of various kinds are inscribed in – and to a great extent, constituted by – everyday communicative practices and social interactions. For language socialization researchers, this means that it is possible to investigate the ways in which everyday exchanges such as those between a young child and his or her primary caregiver relate to other domains of social structure and cultural meaning such as cosmological belief systems, kinship, and patterns of exchange and reciprocity.

The cultural knowledge that guides an individual's participation in mundane social interactions tends to be difficult for the anthropologist or other investigator to tap into because it is implicit, common-sense knowledge that is seldom reflected upon or articulated. But even adults or experts for whom it is second nature realize that it does not just come naturally to the child or novice. Virtually all socialization activities (even those that do not involve explicit teaching) are routinized to some degree, so as to provide the child or novice with opportunities to engage in more or less predictable, schematically structured interactions with caregivers, teachers, or other more experienced persons. Socialization activities are thus contexts in which the background knowledge that adults or experts draw upon during the course of everyday activities tends to be discursively formulated and explicitly articulated for the benefit of children or novices. To be sure, the routinized, explicit qualities of socialization activities also make them a prime analytic focus for the investigator seeking to understand the underlying cultural principles that organize day-to-day social life in the community as a whole.

Transformation and Change

Language socialization research is not merely a matter of accounting for linguistic and cultural reproduction and continuity, however. Language socialization researchers recognize that everyday communicative practices are finely guided by preferences, orientations, and dispositions that are social in origin and culturally specific in nature. But at the same time, they are creatively and strategically deployed by individuals whose particular configurations of interests, intentions, and goals are uniquely their own. Language socialization research therefore highlights the open-ended, negotiated, sometimes contested nature of everyday life, and recognizes that the most ordinary activities may be sites of innovation and far-reaching change.

Languages themselves change over time, and language is a crucial medium through which virtually any aspect of culture may be contested, resisted, renegotiated, and ultimately transformed. Recognizing this, language socialization researchers regard both language and culture as fundamentally emergent, dynamic, open-ended domains that partake of all of the ebbs and flows of social life. A corollary of this is that any given language or culture greatly exceeds the capacities (cognitive, practical, and otherwise) of any individual speaker or social actor; every individual's access, therefore, is necessarily partial (Schieffelin, 1990). This partiality is contingent on specific configurations of social variables (gender, birth order, class, ethnicity, etc.) that in turn shape, and are shaped by, the individual's lived experiences. The variation among individuals that inevitably results – even among children in the same household, or students in the same classroom cohort – is a primordial source of social dynamics, which over the course of time give rise to innovation and change in language and culture alike.

Toward a Relational, Nonlinear Perspective on Human Development

Language socialization researchers therefore avoid conceptualizing human development as a matter of individuals acquiring pre-existent bodies of cultural knowledge, linguistic forms, etc., that are static and bounded. Rather, language and culture, and individuals' acquisition thereof, are conceptualized in relational terms, emphasizing their symbolically mediated, coconstructed, dynamically emergent qualities (Kramsch, 2002).

This is not to say that the heuristic value of established notions of system and structure are spurned in favor of a radical antiformalism. In practical terms, the processes with which language socialization research is centrally concerned can hardly be described and analyzed without reference to some such notions (e.g., grammar, code, speech community), nor without due regard for the demonstrable formal and structural properties that make them coherent units of analysis. The challenge is to avoid unduly reifying such systems and structures: to avoid treating them as self-contained, autonomous entities that somehow exist apart from the historically particular social worlds and lived experiences out of which they emerge, and by virtue of which they can be discerned, labeled, and posited as units of analysis in the first place. The grammar of a given language, for example, is better regarded as a **precipitate of** or **distillation** from the ongoing flow and flux of communicative

practice (at a specific historical moment in the social lives of actual speakers) than as an abstract, timeless **basis for** or **prerequisite of** communicative practice (by hypothetical speakers within an idealized, homogeneous community).

Similar concerns inform language socialization researchers' nonteleological perspectives on the outcomes of socialization (Kramsch, 2002). Individual development is recognized to be variable, contingent, nonlinear, and ultimately open-ended. Differing degrees and types of developmental progress and multiple kinds of successful outcomes are recognized – even when the arena of investigation is a classroom or other institutional context in which the participants themselves differentiate success and failure in much starker terms.

The relational tensions that language socialization researchers seek to explore are well captured by Bourdieu's notion of 'habitus' (Bourdieu, 1977; Bourdieu and Wacquant, 1992). Habitus can be defined as an integrated set of durable, embodied dispositions which predispose the individual to act and react in ways that are more or less specifiable, more or less predictable, but ultimately open-ended and underdetermined. Habitus is the emergent outcome of numerous densely intersecting factors – virtually all of them mediated and substantially influenced by culture – ranging from gender and class status to the particulars of individual life experience. An individual's habitus predisposes him or her to perceive, think, and act in semi-routinized ways, and to regard certain conditions in the world around him or her as normal. But habitus does not rigidly determine that individual's behavior; it is open-ended, allowing for creativity, improvisation, and innovation.

A defining characteristic of habitus is that it is inculcated, i.e., socialized; the dispositions to which Bourdieu refers are said to be acquired largely through socialization, particularly during the early years of the lifespan. Bourdieu's work says very little about how socialization actually comes about, however. Language socialization research strives to fill this gap in our knowledge of human development by means of empirically grounded studies, and in so doing, to contribute to understandings of the multiplex relationships among language, society, and culture.

See also: Anthropological Linguistics: Overview; Bilingual Language Development: Early Years; Communicative Competence; Communities of Practice; Cross-Linguistic Comparative Approaches to Language Acquisition; Early Socialization and Language; Identity in Sociocultural Anthropology and Language; Language Acquisition Research Methods; Language Change and Cultural Change; Language Development: Overview; Language Ideology; Linguistic Anthropology; Social Aspects of Pragmatics; Socialization; Speech and Language Community.

Bibliography

Baquedano-López P (2001). 'Creating social identities through *doctrina* narratives.' In Duranti A (ed.) *Linguistic anthropology: a reader.* Malden, MA: Blackwell. 343–358.

Bloom L (1998). 'Language acquisition in its developmental context.' In Damon W *et al.* (eds.) *Handbook of child psychology,* vol. 2. New York: John Wiley & Sons. 309–370.

Blum-Kulka S & Snow C (eds.) (2002). *Talking to adults: the contribution of multiparty discourse to language acquisition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Bugental D B & Goodnow J J (1998). 'Socialization processes.' In Damon W & Eisenberg N (eds.) *Handbook of child psychology,* vol. 3. New York: John Wiley & Sons. 389–462.

Bourdieu P (1977). *Outline of a theory of practice.* Cambridge: Cambridge University Press.

Bourdieu P & Wacquant L J D (1992). *An invitation to reflexive sociology.* Chicago: University of Chicago Press.

Capps L & Ochs E (1995). *Constructing panic: the discourse of agoraphobia.* Cambridge, MA: Harvard University Press.

Clancy P M (1999). 'The socialization of affect in Japanese mother-child conversation.' *Journal of Pragmatics 31(11),* 1397–1421.

de León L (1998). 'The emergent participant: interactive patterns in the socialization of Tzotzil (Mayan) infants.' *Journal of Linguistic Anthropology 8(2),* 131–161.

Fader A (2001). 'Literacy, bilingualism, and gender in a Hasidic community.' *Linguistics and Education 12(3),* 261–283.

Field M (2001). 'Triadic directives in Navajo language socialization.' *Language in Society 30(2),* 249–263.

Garrett P B (2005). 'What a language is good for: language socialization, language shift, and the persistence of code-specific genres in St. Lucia.' *Language in Society 34(3).*

Garrett P B & Baquedano-López P (2002). 'Language socialization: reproduction and continuity, transformation and change.' *Annual Review of Anthropology 31,* 339–361.

Gutiérrez K, Baquedano-López P & Alvarez H (2001). 'Using hybridity to build literacy in urban classrooms.' In Reyes M L & Halcón J J (eds.) *The best for our children: Latina/Latino voices in literacy.* New York: Teachers College Press. 122–141.

Heath S B (1983). *Ways with words: language, life, and work in communities and classrooms.* Cambridge: Cambridge University Press.

Kramsch C (ed.) (2002). *Language acquisition and language socialization: ecological perspectives.* New York: Continuum.

Kulick D (1992). *Language shift and cultural reproduction: socialization, self, and syncretism in a Papua New Guinean village*. Cambridge: Cambridge University Press.

Kulick D & Schieffelin B B (2004). 'Language socialization.' In Duranti A (ed.) *A companion to linguistic anthropology.* Malden, MA: Blackwell. 349–368.

Moore L C (2004). 'Multiligualism and second language acquisition in the Northern Mandara Mountains of Cameroon.' In Echu G & Obeng S G (eds.) *Africa meets Europe: language contact in West Africa.* New York: Nova Science Publishers.

Ochs E (1988). *Culture and language development: language acquisition and language socialization in a Samoan village.* Cambridge: Cambridge University Press.

Ochs E & Schieffelin B B (1984). 'Language acquisition and socialization: three developmental stories and their implications.' In Shweder R A & LeVine R A (eds.) *Culture theory: essays in mind, self and emotion.* New York: Cambridge University Press. 276–320.

Ochs E & Schieffelin B B (1995). 'The impact of language socialization on grammatical development.' In Fletcher P & MacWhinney B (eds.) *The handbook of child language*. Cambridge, MA: Basil Blackwell. 73–94.

Ochs E, Smith R & Taylor C (1989). 'Detective stories at dinnertime: problem-solving through co-narration.' *Cultural Dynamics 2,* 238–257.

Ochs E, Taylor C, Rudolph D & Smith R (1992). 'Storytelling as a theory-building activity.' *Discourse Processes 15(1),* 37–72.

Paugh A (2005). 'Multilingual play: children's code-switching, role play, and agency in Dominica, West Indies.' *Language in Society 34(1),* 63–86.

Platt M (1986). 'Social norms and lexical acquisition: a study of deictic verbs in Samoan child language.' In Schieffelin B B & Ochs E (eds.) *Language socialization across cultures.* Cambridge: Cambridge University Press. 127–152.

Schieffelin B B (1990). *The give and take of everyday life: language socialization of Kaluli children.* Cambridge: Cambridge University Press.

Schieffelin B B (1996). 'Creating evidence: making sense of written words in Bosavi.' In Ochs E, Schegloff E A & Thompson S A (eds.) *Interaction and grammar.* Cambridge: Cambridge University Press. 435–460.

Schieffelin B B & Ochs E (1986a). 'Language socialization.' *Annual Review of Anthropology 15,* 163–191.

Schieffelin B B & Ochs E (eds.) (1986b). *Language socialization across cultures.* Cambridge: Cambridge University Press.

Schieffelin Bambi B & Ochs Elinor (1996). 'The microgenesis of competence: methodology in language socialization.' In Slobin D I, Gerhardt J, Kyratzis A & Guo J (eds.) *Social interaction, social context, and language: essays in honor of Susan Ervin-Tripp*. Mahwah, NJ: Lawrence Erlbaum Associates. 251–264.

Watson-Gegeo K A (1992). 'Thick explanation in the ethnographic study of child socialization: a longitudinal study of the problem of schooling for Kwara'ae (Solomon Islands) children.' *New Directions for Child Development 58,* 51–66.

Watson-Gegeo K A & Nielsen S (2003). 'Language socialization in SLA.' In Doughty C J & Long M H (eds.) *The handbook of second language acquisition.* Malden, MA: Blackwell. 155–177.

Language Spread

S S Mufwene, University of Chicago, Chicago, IL, USA

The term 'language spread' alternates in the literature with such other terms as 'language expansion,' 'language dispersal,' and 'language diffusion' in reference to the fact that a language acquires more speakers who use it as a vernacular or as a lingua franca. The process may or may not be a concomitant of geographical expansion due, for instance, to trade or to the colonization of other territories by its speakers. A critical factor is its adoption by another group, not the demographic increase of the original population. This explanation applies as much to the fact that English has been gaining more and more speakers around the world as to the fact that it gradually acquired more and more speakers among the Celts in the British Isles, having originated between the 6th and 7th centuries as the language of the Germanic invaders. It is also as true of the spread of Latin from Rome to the surrounding areas, in Italy, in the 4th century B.C. as of its dispersal in the Roman Empire, where it transformed into the Romance languages. It likewise applies to the gradual Romanization of the Romance countries, at the expense of the more indigenous Celtic languages, as well as to the dispersal of some of the Romance languages (French, Portuguese, and Spanish) around the world and their transformation in some places (e.g., the Caribbean and Indian Ocean former plantation settlement colonies) into vernaculars now disfranchised as Creoles. The evolution is just as true of Swahili too, in East Africa, bearing in mind that it was so named by the Arab traders in the 7th century in association with the coastal populations who spoke related Bantu languages from Kenya to Mozambique. They used its

Kulick D (1992). *Language shift and cultural reproduction: socialization, self, and syncretism in a Papua New Guinean village.* Cambridge: Cambridge University Press.

Kulick D & Schieffelin B B (2004). 'Language socialization.' In Duranti A (ed.) *A companion to linguistic anthropology.* Malden, MA: Blackwell. 349–368.

Moore L C (2004). 'Multiligualism and second language acquisition in the Northern Mandara Mountains of Cameroon.' In Echu G & Obeng S G (eds.) *Africa meets Europe: language contact in West Africa.* New York: Nova Science Publishers.

Ochs E (1988). *Culture and language development: language acquisition and language socialization in a Samoan village.* Cambridge: Cambridge University Press.

Ochs E & Schieffelin B B (1984). 'Language acquisition and socialization: three developmental stories and their implications.' In Shweder R A & LeVine R A (eds.) *Culture theory: essays in mind, self and emotion.* New York: Cambridge University Press. 276–320.

Ochs E & Schieffelin B B (1995). 'The impact of language socialization on grammatical development.' In Fletcher P & MacWhinney B (eds.) *The handbook of child language.* Cambridge, MA: Basil Blackwell. 73–94.

Ochs E, Smith R & Taylor C (1989). 'Detective stories at dinnertime: problem-solving through co-narration.' *Cultural Dynamics 2,* 238–257.

Ochs E, Taylor C, Rudolph D & Smith R (1992). 'Storytelling as a theory-building activity.' *Discourse Processes 15(1),* 37–72.

Paugh A (2005). 'Multilingual play: children's code-switching, role play, and agency in Dominica, West Indies.' *Language in Society 34(1),* 63–86.

Platt M (1986). 'Social norms and lexical acquisition: a study of deictic verbs in Samoan child language.' In Schieffelin B B & Ochs E (eds.) *Language socialization across cultures.* Cambridge: Cambridge University Press. 127–152.

Schieffelin B B (1990). *The give and take of everyday life: language socialization of Kaluli children.* Cambridge: Cambridge University Press.

Schieffelin B B (1996). 'Creating evidence: making sense of written words in Bosavi.' In Ochs E, Schegloff E A & Thompson S A (eds.) *Interaction and grammar.* Cambridge: Cambridge University Press. 435–460.

Schieffelin B B & Ochs E (1986a). 'Language socialization.' *Annual Review of Anthropology 15,* 163–191.

Schieffelin B B & Ochs E (eds.) (1986b). *Language socialization across cultures.* Cambridge: Cambridge University Press.

Schieffelin Bambi B & Ochs Elinor (1996). 'The microgenesis of competence: methodology in language socialization.' In Slobin D I, Gerhardt J, Kyratzis A & Guo J (eds.) *Social interaction, social context, and language: essays in honor of Susan Ervin-Tripp.* Mahwah, NJ: Lawrence Erlbaum Associates. 251–264.

Watson-Gegeo K A (1992). 'Thick explanation in the ethnographic study of child socialization: a longitudinal study of the problem of schooling for Kwara'ae (Solomon Islands) children.' *New Directions for Child Development 58,* 51–66.

Watson-Gegeo K A & Nielsen S (2003). 'Language socialization in SLA.' In Doughty C J & Long M H (eds.) *The handbook of second language acquisition.* Malden, MA: Blackwell. 155–177.

Language Spread

S S Mufwene, University of Chicago, Chicago, IL, USA

The term 'language spread' alternates in the literature with such other terms as 'language expansion,' 'language dispersal,' and 'language diffusion' in reference to the fact that a language acquires more speakers who use it as a vernacular or as a lingua franca. The process may or may not be a concomitant of geographical expansion due, for instance, to trade or to the colonization of other territories by its speakers. A critical factor is its adoption by another group, not the demographic increase of the original population. This explanation applies as much to the fact that English has been gaining more and more speakers around the world as to the fact that it gradually acquired more and more speakers among the Celts in the British Isles, having originated between the 6th and 7th centuries as the language of the Germanic invaders. It is also as true of the spread of Latin from Rome to the surrounding areas, in Italy, in the 4th century B.C. as of its dispersal in the Roman Empire, where it transformed into the Romance languages. It likewise applies to the gradual Romanization of the Romance countries, at the expense of the more indigenous Celtic languages, as well as to the dispersal of some of the Romance languages (French, Portuguese, and Spanish) around the world and their transformation in some places (e.g., the Caribbean and Indian Ocean former plantation settlement colonies) into vernaculars now disfranchised as Creoles. The evolution is just as true of Swahili too, in East Africa, bearing in mind that it was so named by the Arab traders in the 7th century in association with the coastal populations who spoke related Bantu languages from Kenya to Mozambique. They used its

Zanzibari variety as a trade language while penetrating the African interior, an example that the European colonizers (the Germans, the English, and the Belgians) would follow in adopting it in the East African region (from the coast to eastern Congo) for communication with the masses of the populations.

The causes of language spread are diverse, although they boil down to population contact, obviously under a variety of interactional conditions. We do not know some of the significant reasons, though linguists must figure out some way of articulating 'costs' and 'benefits' to those who adopt it. Power and prestige have often been invoked as the reason. This factor applies in the case of Latin in the Roman Empire, where the local nobilities were first to adopt it, in order to partake in the Roman rule of their own territories. However, the fact that the Roman elite in the eastern Roman Empire preferred a more prestigious Greek language did not favor the latter for long. There was no linguistic Hellenization process comparable to the Latinization of southwestern continental Europe.

It is not clear either why the Arabs and, later, the Europeans adopted indigenous languages, instead of imposing their own (as in most settlement colonies) to communicate with the indigenous populations in, respectively their trade and their exploitation colonies. One can invoke courtesy to local populations they traded with in the case of the Arabs, but the emergence of pidgins based on European languages on the West African coast and in the Pacific shows that early on, before the exploitation colonization of these regions, the Europeans preferred to use their own languages. Even the latter generalization is not so accurate, because trade pidgins in the Americas were based on Native American languages, while European languages were being adopted as vernaculars in the settlement colonies. Trade with Native Americans produced new varieties such as Delaware Pidgin in the New Netherland, the Mobilian Jargon (Mobilian) in the southeastern United States, and Lingua Geral (Nhengatu) in Brazil; whereas plantation settlement colonies produced creoles such as Haitian (Haitian Creole French), Jamaican (Creole English), and Papiamentu and other colonies produced noncreole varieties such as American English, Québécois French, and Popular Brazilian Portuguese.

Noteworthy also is the fact that European languages have been retained as official languages in former European exploitation colonies of Africa and Asia, in part because the selection of an indigenous lingua franca for the same function is feared to arouse political problems of national unity versus ethnic identities and economic problems having to do with the development of tools for teaching it in school. Another reason seems to lie simply in a legacy of the colonial period, viz., the association of the European language with power and prestige. This accounts for why European languages continue to function as official languages in practically monolingual countries such as Rwanda, Burundi, and Botswana. Ability to communicate with the outside world has often been invoked as an important reason, though it is a dubious one even in these days when globalization raises fears of (continued) economic marginalization of economically less developed nations.

Language spread has its structural costs. The prevailing language typically undergoes a Pyrrhic victory, changing into a new variety within the new community, regardless of whether it is acknowledged as the same language or a different one. Regardless of whether Latin can be claimed to still be alive and spoken in various places around the world, its new forms are identified by different names, e.g., French, Spanish, Portuguese, Haitian Creole, Papiamentu, or Cape Verdian Criolou (Kabuverdianu). The reasons for these name changes are as political as they are structural. In some cases, metropolitan speakers deny that particular colonial varieties are still the same language. Most significant is the fact that populations appropriating a foreign language influence its evolution with features from their substrate languages.

The case of English is particularly interesting to discuss, because its origins have received less attention than they deserve, though they should inspire research on the beginnings of other (major) languages. The original Germanic populations that colonized England in the 5th century (viz., the Jutes, the Angles, the Saxons) brought no English language with them. They spoke Germanic languages that may even not have been mutually intelligible. (One of them may be related to the present Jutlandish dialect of Danish.) While it can be argued that none of these languages prevailed, or maybe that it was the language of the Angles (who have bequeathed their names to both the language and the territory now identified as *England*) that did, Old English appears to have emerged from contacts primarily from these immigrants' languages. Successive transformations of the latter into varieties called Middle and Modern English, and later into, concurrently, indigenized Englishes (e.g., Indian and West African English), pidgin English (e.g., Nigerian Pidgin English [Nigerian Pidgin] and Tok Pisin in Papua New Guinea), and creole English (e.g., Jamaican and Hawaiian [Hawai'i Creole English] Creole English) all reflect structural changes that the language has undergone at different stages of its evolution under the influence of other languages it came in contact with.

An important question remains whether different heuristic approaches are needed (as is the dominant trend in the linguistics literature to date) to account for this differential evolution of a language. Both Bantu and Indo-European languages seem to have spread and diversified due to influence from languages spoken by the populations that their speakers colonized, although the contact histories are far from reflecting the oversimplified scenarios suggested by the cladograms or Stammbaums traditionally used to illustrate genetic ties among languages. The diversification of Arabic can likewise be associated with substrate influence as a factor bearing on the evolution of language spreading among new populations.

It must also be borne in mind that languages do not move about on their own, nor do they compete directly with each other, contrary to what is suggested by the growing literature on language endangerment and loss, in which languages are misportrayed as warring with and killing each other. As a matter of fact languages typically spread faster in times of peace, when individuals from separate ethnic groups or nationalities are more likely to interact with each other. Languages are analogous to parasitic species in biology; their lives depend largely on the social practices of their hosts and makers, i.e., their speakers. Moreover, critical agents in their spread need not be their native speakers.

Recent colonial and postcolonial history suggests that nonnative speakers participate in the process too. The literature on indigenized Englishes makes it quite obvious that English has spread significantly after the independence of Great Britain's former exploitation colonies and its adoption as those colonies' official language, with the language being taught more and more by nonnative speakers. Where teachers are native speakers, such as in the West Indian former plantation settlement colonies, they do not speak the metropolitan varieties of the United Kingdom and North America and therefore contribute to the ongoing diversification of the common language of the former British Empire. A large proportion of teachers of English in several other countries that since then have adopted it as an international language come from the same territories where English is nonmetropolitan. Note that even in places where English functions as a native vernacular of the majority, several immigrants learn English naturalistically more from other nonnative speakers, with whom they socialize more regularly, than from native speakers. *Mutatis mutandis*, this evolutionary pattern applies to other European colonial languages of the last half-millennium.

A reexamination of the history of the Roman Empire also suggests an important contribution of nonnative speakers to the spread and divergence of Vulgar Latin into the diverse Romance languages. In addition to the fact that Latin spread among the masses of the former colonies mostly after the Romans had left their western empire, a significant proportion of model speakers of Vulgar Latin during the Roman rule consisted of legionaries who served as mercenaries or had been recruited from the local populations and were nonnative speakers. It can be conjectured that similar evolutions took place under similar conditions of colonial or imperial expansion, factoring in the role of interpreters in the establishment of initial intergroup communication.

How interpreters learned the foreign language, if it was not spoken by a neighboring group, deserves investigating and certainly more attention than their role has received. Recent European colonial history shows that they played an important role in China, on the West African coast, and in Hawaii, leading to the development of pidgin varieties of European languages. It is not the interpreters as linguistic experts who created these pidgin varieties. It is the masses of the local populations, who, as trade with Europeans increased, would interact directly with Europeans using a language that they definitely commanded less well than the interpreters. (The proportion of the latter decreased as the trade expanded.) The contributions of Europeans in responding in foreigner talk or, in some cases, speaking the European trade language no better than the natives (e.g., French or Dutch traders using Portuguese on the West African coast) should not be underestimated.

Much of the literature has underscored the role of schools and the media in the spread of languages. Indeed this factor cannot be underrated in modern societies, in which all sorts of techniques have been developed for efficient language teaching. However, we must recall that spoken languages antedate written ones and even before the invention of writing systems languages used to spread too. To date there are also many societies which have large proportions of their populations which are multilingual but functionally illiterate. They have learned languages other than their native ones simply through interactions with the latter's speakers and by oral communication, without the mediation of schools. As a matter of fact, even in predominantly literate societies, an important factor in language spread is that learners must have other individuals with whom they can practice the language. Media without active interaction provide only passive knowledge of the relevant language. European languages have continued to spread in former exploitation colonies thanks largely to former colonial auxiliaries who at the wake of independence became the new political leaders and continued to use

the colonial language in their official functions and the school system. The association of these languages with indigenous populations and specific local ethnographic functions within the new nation-states, as opposed to their former association with foreign colonial rule, enhanced their usefulness to the masses of the population. Many of the features now characterized as substrate in extra-European varieties reflect that cumulative mutual reinforcement of new communicative habits within the local populations. The local or national media only reinforce them.

The view of language spread presented here underscores the important role of language contact as a factor affecting language spread. It bears significantly on how we must deal with language diversification in genetic linguistics and historical dialectology. It questions the 'asexual,' monoparental tradition of the comparative method and the summary of its findings in Stammbaums without factoring in the influences that various languages have exerted on each other's evolution, leading sometimes to the formation of 'linguistic areas.'

See also: Endangered Languages; English: World Englishes; Language Change and Language Contact; Language Maintenance and Shift; Languages of Wider Communication; Migration and Language; Pidgins and Creoles: Overview.

Bibliography

Aikhenvald A Y & Dixon R M W (eds.) (2001). *Areal diffusion and genetic inheritance: problems in comparative linguistics.* Oxford: Oxford University Press.

Cavalli-Sforza L L (2000). *Genes, peoples, and languages.* New York: North Point Press.

Cooper R L (ed.) (1982). *Language spread.* Bloomington: Indiana University Press.

Martinet A (1986). *Des steppes aux océans: l'indo-européen et les 'Indo-européens.'* Paris: Payot.

Mazrui A & Mazrui A (1998). *The power of Babel: language in the African experience.* Oxford: James Currey/ Chicago: University of Chicago Press.

Mufwene S S (2001). *The ecology of language evolution.* Cambridge: Cambridge University Press.

Nettle D (1999). *Linguistic diversity.* Oxford: Oxford University Press.

Nichols J (1992). *Linguistic diversity in space and time.* Chicago: University of Chicago Press.

Renfrew C (1987). *Archaeology and language: the puzzle of Indo-European origins.* London: Jonathan Cape.

Renfrew C (1994). 'World linguistic diversity.' *Scientific American,* January, 116–123.

Ruhlen M (1994). *The origin of language: tracing the evolution of the mother tongue.* New York: John Wiley & Sons.

Skuttnab-Kangas T (2000). *Linguistic genocide in education – or world-wide diversity and human rights?* Mahwah, NJ: Lawrence Erlbaum Associates.

Language Surveys

J de Vries, Carleton University, Ottawa, Canada

This article discusses the wording of language questions in national censuses and language surveys and what they reveal about the assumptions of governments and other agencies. There is a small and not easily accessible literature on this topic. While most of the material has been contributed by demographers, some crucial entries have been produced by linguists with an interest in the quantitative aspects of such issues as language contact, bilingualism, language shift, and maintenance. The first part of this discussion deals with the modal question types, their meaning and interpretation, and some aspects of data quality. That discussion shows a strong emphasis on the situation in Canada, in part because the best sources are there. Nevertheless, there are some examples from other countries, without any claim of completeness. The second part of this article is a brief discussion of language surveys that are designed to solicit information about the language characteristics of individuals.

Language Questions in Censuses

Language questions in censuses have been asked in many countries for a long time. According to Arel, the first instances were in Belgium and Switzerland (2002), though no years were given in that source. Information obtained from the Swiss Bureau of Federal Statistics shows 1860 as the earliest appearance of a question on language. The earliest Belgian census to contain a language question was in 1846. Arel provided a broader historical overview and an interesting discussion of the international debates on the utility of census questions on language, ethnicity, and nationality.

For most countries, the inclusion of language questions in censuses is mandated in order to administer particular policies, such as providing information

the colonial language in their official functions and the school system. The association of these languages with indigenous populations and specific local ethnographic functions within the new nation-states, as opposed to their former association with foreign colonial rule, enhanced their usefulness to the masses of the population. Many of the features now characterized as substrate in extra-European varieties reflect that cumulative mutual reinforcement of new communicative habits within the local populations. The local or national media only reinforce them.

The view of language spread presented here underscores the important role of language contact as a factor affecting language spread. It bears significantly on how we must deal with language diversification in genetic linguistics and historical dialectology. It questions the 'asexual,' monoparental tradition of the comparative method and the summary of its findings in Stammbaums without factoring in the influences that various languages have exerted on each other's evolution, leading sometimes to the formation of 'linguistic areas.'

See also: Endangered Languages; English: World Englishes; Language Change and Language Contact; Language Maintenance and Shift; Languages of Wider Communication; Migration and Language; Pidgins and Creoles: Overview.

Bibliography

Aikhenvald A Y & Dixon R M W (eds.) (2001). *Areal diffusion and genetic inheritance: problems in comparative linguistics.* Oxford: Oxford University Press.
Cavalli-Sforza L L (2000). *Genes, peoples, and languages.* New York: North Point Press.
Cooper R L (ed.) (1982). *Language spread.* Bloomington: Indiana University Press.
Martinet A (1986). *Des steppes aux océans: l'indo-européen et les 'Indo-européens.'* Paris: Payot.
Mazrui A & Mazrui A (1998). *The power of Babel: language in the African experience.* Oxford: James Currey/Chicago: University of Chicago Press.
Mufwene S S (2001). *The ecology of language evolution.* Cambridge: Cambridge University Press.
Nettle D (1999). *Linguistic diversity.* Oxford: Oxford University Press.
Nichols J (1992). *Linguistic diversity in space and time.* Chicago: University of Chicago Press.
Renfrew C (1987). *Archaeology and language: the puzzle of Indo-European origins.* London: Jonathan Cape.
Renfrew C (1994). 'World linguistic diversity.' *Scientific American,* January, 116–123.
Ruhlen M (1994). *The origin of language: tracing the evolution of the mother tongue.* New York: John Wiley & Sons.
Skuttnab-Kangas T (2000). *Linguistic genocide in education – or world-wide diversity and human rights?* Mahwah, NJ: Lawrence Erlbaum Associates.

Language Surveys

J de Vries, Carleton University, Ottawa, Canada

This article discusses the wording of language questions in national censuses and language surveys and what they reveal about the assumptions of governments and other agencies. There is a small and not easily accessible literature on this topic. While most of the material has been contributed by demographers, some crucial entries have been produced by linguists with an interest in the quantitative aspects of such issues as language contact, bilingualism, language shift, and maintenance. The first part of this discussion deals with the modal question types, their meaning and interpretation, and some aspects of data quality. That discussion shows a strong emphasis on the situation in Canada, in part because the best sources are there. Nevertheless, there are some examples from other countries, without any claim of completeness. The second part of this article is a brief discussion of language surveys that are designed to solicit information about the language characteristics of individuals.

Language Questions in Censuses

Language questions in censuses have been asked in many countries for a long time. According to Arel, the first instances were in Belgium and Switzerland (2002), though no years were given in that source. Information obtained from the Swiss Bureau of Federal Statistics shows 1860 as the earliest appearance of a question on language. The earliest Belgian census to contain a language question was in 1846. Arel provided a broader historical overview and an interesting discussion of the international debates on the utility of census questions on language, ethnicity, and nationality.

For most countries, the inclusion of language questions in censuses is mandated in order to administer particular policies, such as providing information

about national elections (for the United States, see Jiu, 1999 and Crawford, 2002) or access to public primary or secondary education in a minority language or for special education in an official language. (For Canada, see Martel, 1991. For the United States, see Crawford, 2002; Stevens, 1999). It should be recognized that language questions in censuses are not designed to measure fine-grained linguistic features. In addition, there are several general methodological problems. Data such as these are collected either in interviews between a census employee (the enumerator) and a respondent, or else they are self-completed written questionnaires. The latter method has become the most common one in most developed countries, though electronic media may eventually become the method of choice. Completion of the census over the internet has already been possible in, among other countries, Switzerland, Spain, and Singapore; in 2006 Canada can be added to that list. Census enumerators are not thoroughly trained in interviewing techniques, and in many cases the self-response modes provide the respondent with only minimal instructions, at best. Another complication is that census data collection methods generally involve responses given by one household member "on behalf of" the other members (the so-called proxy response). Finally, censuses are rather expensive tools for collecting any social data, including linguistic information. Regardless of the data collection mode selected, the instrument is, by necessity, rather brief. It is rare to find more than one question about language in a census, let alone the wide array of questions necessary for a researcher to get a clear picture of the prevalence of bilingualism or the fluency of different population groups in a particular language. In summary, language questions in censuses should be considered only as indicators of underlying linguistic concepts. The fact that they all involve verbal responses by members of the public implies that such indicators may have less than perfect validity and reliability.

In defense of these data, it should be noted that the few studies aimed at validating such self-response data have, overall, shown that validity tends to be satisfactory (on the condition that the social situation is conducive to unbiased responding) and that reliability likewise is within the acceptable range. Linguists' arguments against using census data are therefore not tenable, as long as these reservations are kept in mind. Obviously, finer-grained linguistic issues (e.g., about size of lexicon, or spoken accent, or the use of proper grammar) cannot be addressed with these simple indicators, but they are suitable for the broader applications for which these questions are mandated. As with most data collected for social research and/or policy, the motto should be '*caveat emptor!*'

Over the past eighty years or so, several authors have written articles on this topic. This paper may be seen as an update on those earlier publications (see, for example, Kloss, 1929; Lieberson, 1966, 1969; de Vries, 1985, 1990, 2004, to mention only a few).

Types of Questions

Census questions on language may be grouped into three categories.

Mother Tongue The usual operational definition of questions regarding the mother tongue is 'the language learned in earliest childhood.' Such questions are assumed to indicate the linguistic origins of a person. In the case of minority language education in Canada (essentially, French outside Quebec, English in Quebec), the mother tongue of one of the parents determines the eligibility of a child to obtain publicly funded education in the regional minority language. Similarly, in Finland, the mother tongue composition of the population of a municipality or province determines whether or not public services are provided in the minority language (usually Swedish). It should be noted that several countries have deviant specifications. In the censuses of Switzerland from 1860 until 1980, for example, the question wording, in each of the official languages of the country, refers to a "main language," despite being labeled as a "mother tongue" question. The Indian census of 2001 literally asked for the language spoken to the respondent by his or her mother. The most recent Canadian version specifies "the language first learned at home in childhood and still understood." It should be noted that the final specification in the Canadian definition allows for a person to have lost his or her mother tongue. Among immigrants to Canada, such a possibility is not really far-fetched. The Canadian census instructions tell persons who are no longer able to understand the first language learned to supply the next language, that is, the language they acquired next.

Bondarskaya (1993), in a general discussion of Russian census data on nationality, claimed that "information on ethnic origin, mother tongue and religious denomination was gathered only during the first five censuses, the first conducted in 1719, and the fifth in 1795." The reference to a census as early as 1719 is somewhat unusual, since the general demographic literature does not refer to such early censuses in Russia. However, it is indisputable that the Soviet empire did conduct many censuses, beginning in 1920. Many of those did contain a question on mother tongue. The annotations in the appendix to

Bondarskaya's paper made it clear that these questions were generally not proper mother tongue questions. In 1920 the question referred to the "language spoken by the family of the respondent (or by the mother in multilingual families)." In 1926 the mother tongue was defined as "the language in which the respondent is most fluent or usually speaks." The most recent censuses discussed by Bondarskaya (those of 1970, 1979, and 1989) had consistent definitions of the mother tongue: "the language considered as native by the respondent himself is given as the mother tongue. If the respondent has difficulty in deciding on his mother tongue, the language in which the respondent is most fluent or which he usually speaks in the family should be given as his mother tongue" (1993). Note the contrast with the Canadian case, where respondents are asked to report the "next language learned" if they no longer understand the first one acquired, and to select the earlier one of two (or more) if they acquired more than one language in early childhood. In the Canadian censuses, the issue of fluency only emerges for those respondents who not only acquired two languages in early childhood but who, in addition, cannot (or will not) determine which one was acquired first.

Main Language Questions of addressing a respondent's main language, or some functional equivalent, address the issue of **use**. In contrast to the mother tongue type of questions (which refers to a person's past) questions of the main language pertain to the present. Some versions ask for the language(s) spoken most often in a specific domain, such as the respondent's home (Canada), work or school (the censuses for Switzerland in 1990 and 2000). I am not sure whether any country is currently using the "main language" indicator as an instrument for public policy. Until the language question was changed, the classification of administrative units in Finland (municipalities, provinces) was based on a main language classification (Finnäs, 1986).

The Canadian census question in this category asked, "What language does this person speak most often at home?" Pre-coded response categories were English and French. Other languages were captured as "Other (specify)" with a write-in box. Researchers interested in trend analyses on data such as these should be aware that question wording as well as response categories changed several times since the first appearance of this question in 1971.

Ability to Speak Designated (Official) Languages Another type of questions addresses the ability to speak the official languages of a country (English or French in Canada) or a region (Welsh in Wales, Gaelic in parts of Scotland). They address **capacity**, rather than history or current use. Their function is to assess whether there is a sufficiently large number of persons able, or unable, to speak an official language, to warrant public services in a (regional) minority language. Both the United States and Australia asked a question regarding a person's ability to speak English (Crawford, 2002; Clyne, 1991). The Canadian question in this category asked, "Can this person speak English or French well enough to conduct a conversation?" The response categories were: English only; French only; both English and French; neither English nor French.

Bondarskaya's discussion informs us that the most recent three Soviet censuses asked, in addition to the mother tongue question, whether the respondent is able to speak "any other language of the U.S.S.R." (1993). In fact, the question appears to refer to fluency. Note the restriction to the languages of the Soviet Union. If a respondent wrote in more than one of these languages, only the first one mentioned was processed; subsequent responses were ignored. Note also that this question did not pertain to fluency in other languages of wider circulation, such as English, French, or German.

Ability to Speak Other Languages, Not an Official Language Similar to the questions on "ability in official languages," another type of question refers, in its strictest form, to **capacity**, although several question wordings stray into the **current use** territory, as with the Canadian census questions from 1991 on the ability to speak non-official languages. The question in the Canadian census of 2001 was a follow-up question to the one on home language, in which a person was asked: "Does this person speak any other languages on a regular basis at home?" The response categories were: *No; Yes, English; Yes, French; Yes, other – specify.* The layout of the questionnaire did not allow for multiple responses, though such a situation would not have been impossible.

Del Pinal and Lapham (1993) discussed the U.S. census question on language for 1990. This was a three-part question, beginning with "Does this person speak a language other than English at home?" with response categories *Yes* or *No*. The instructions indicate that *Yes* should be marked if the respondent "sometimes or always" speaks a language other than English at home. For those who answered *yes*, there were two follow-up questions. First, "What is this language?" with a write-in option (examples given are Chinese, Italian, Spanish, Vietnamese). If the person speaks more than one non-English language at home, the one spoken most often should be recorded; if the person cannot determine which language is

spoken more often, the one first learned by the person should be included. Census processing obviously ignored any mention beyond the first one. Note that for some, but not all, respondents, the language other than English would be identical with the mother tongue. The second follow-up question for those who do speak another language at home asked for their self-reported fluency in English: "How well does this person speak English?" The response categories were: *Very well; Well; Not well; Not at all.*

Del Pinal and Lapham's figure 7 (1993) summarized the various census questions related to ethnicity. The question on non-English language was asked in the censuses of 1890, 1910 through 1930, and 1980 through 1990. In the last two of those censuses the question was only asked in the long form, administered only to a sample of the residential population. The censuses of 1910 through 1940, as well as 1960 and 1970, also asked for the mother tongue of the respondent and his/her foreign-born parents.

As Stevens points out, the discontinuity in question wordings, response categories, and general methodology renders any longitudinal comparisons of these data virtually impossible. Various censuses shifted between an emphasis on the use of languages other than English and a person's fluency in English. "The outcome is an insurmountable barrier to comparison across the censuses in which only a question on non-English language usage was asked and those in which only a question on English proficiency was included" (Stevens, 1999).

General Comments It should be noted that each one of the varieties of language questions has its shortcomings. For Canada we know, for example, that multiple responses to the questions on mother tongue and home language have been given, despite the fact that such multiple responses have always been strongly discouraged by the census designers. This issue is not a new one, nor is it apparently limited to Canada. Lieberson, in his analysis of the measurement of bilingualism, commented: "censuses seem to never recognize that a respondent could have more than one mother tongue" (1969).

Each one of these language questions is subject to potential political or social pressures. It may, for example, be expedient for members of a particular language community to understate their ability to speak the language of the linguistic or political majority. This did occur, for example, in Estonia in 1979, when the number of respondents claiming fluency in Russian declined substantially from that reported in 1970 (Arel, 2002). Another case of invalid responses as a consequence of political pressures occurred in Belgium in the 1947 census, especially in Brussels and its Flemish-speaking suburbs. Since at that time many Flemish-speaking parents were keen on sending their children to a French-speaking school, many of them declared that their main language was French (Arel, 2002). It should be recognized that the potential for deliberate sabotage of the census for political or ideological reasons is not restricted to respondents. As Khubchandani pointed out: "various governments make intentional distortions in taking the censuses of their populations, such as applying varying (inclusive/exclusive) criteria in the grouping of mother tongues into umbrella-like 'languages' or isolating them into distinct entities, as perceived by dominant agencies or pressure groups in a region" (1995).

Methodological Aspects

There is a voluminous literature pertaining to the quality of census data. In general, errors in demographic data are classified into seven broad categories (discussed by Brackstone, 1977).

1. Coverage Errors: A coverage error occur when a person (or, for that matter, anything we want to count) either is not counted (exclusion or undercoverage) or is counted more than once (overcoverage). Another manifestation of overcoverage occurs when someone is counted but should not have been (for example, temporary residents or foreign laborers on work permits). While in most developed societies overall errors of coverage amount to only a small percentage of the total population, differential errors of coverage may have a somewhat larger impact: it may happen that some members of a specific language community are (substantially) more likely to be missed in a census than others, or else have chosen to boycott the census for political or ideological reasons. For example, the 1996 Canadian technical report on coverage showed that net undercoverage (the difference between over- and undercoverage) amounted to 2.18% for persons with English as mother tongue, 1.67% for persons with French as mother tongue, but 5.17% for respondents with mother tongues other than English or French. A reasonable guess is that many of these missing respondents were recent immigrants with limited or no knowledge of the official languages of the country (English or French); however the report does not supply further information on this (Statistics Canada, 1999). At the time of this writing, the technical report on coverage for the 2001 census was not available.
2. Non-Response Errors: A non-response error occurs when a person has been counted, but has not provided responses to all of the questions

which were asked. In survey analysis, these situations are usually classified as "missing data." Below I discuss in greater detail the prevalence of non-response errors and their treatment in the Canadian census of 2001.

3. Response Errors or "Errors of Content": A third type of error is what we call "errors of content." These involve cases where a person or an event has been counted, but where one or more aspects have been measured incorrectly. For example, a person might have responded with the answer "French" to a question on mother tongue, while the correct response was "English." It could also mean that a respondent did not give any answer to a question (a non-response). Finally, there are situations in which respondents provide multiple answers to questions for which only one reply was requested. In such a situation, a person provides a response that is incorrect. An additional cause of response error is due to an interviewer or enumerator misunderstanding a response, thus misrecording what may have been a correct answer.
4. Coding Errors: Censuses, as well as other self-report methodologies, usually involve a stage in which verbal responses are coded, that is, placed into one of a previously defined set of categories. In census processing this occurs whenever responses are given as open-ended answers (e.g., when a respondent has a mother tongue that is not given on the list of provided languages). Typically this involves a prespecified category – "other, not elsewhere specified" – followed by a "please specify" stimulus. Such responses require coding that until recently was performed by humans. The more recent substitution of computer processing has not necessarily eliminated coding error.
5. Data Capture Error: The processing of large volumes of data involves a variety of mechanical processes used to transform the initial responses (usually written forms on paper) into a medium that can be further processed by computers prior to further editing and tabulation.
6. Editing and Imputation Errors: In most cases, census and survey data are subjected to a variety of checks for consistency, non-response error, etc. In many countries such edits and imputations are performed by computers, following a set of rules in which particular inconsistencies are defined and procedures for the resolving them are programmed. There is now a small literature on situations where edit and imputation procedures have, in fact, produced new errors (e.g., by resolving particular inconsistencies in erroneous ways).
7. Sampling Errors: Sampling errors are more common in surveys (which always involve samples), but they also occur in censuses that use more than one form and administer one or more of these alternative forms to a sample of the population. The most common occurrence is where census questionnaires are produced in a long and a short form. The latter is usually given to a small sample of private households – for example, in the 2001 census of population for Canada, 20% of private households were given the long form, while the remaining 80% received the (much shorter) short form. Such sampling has been used in the censuses of almost all developed societies. Whenever samples are taken, the possibility arises that the selected sample is not perfectly representative of the total population. In cases such as these there is a slight possibility that the data provided in the sample differ from those pertaining to the entire population. In the Canadian census data, the question on mother tongue is contained in the short form (due to legal requirements) while the remaining language questions (language spoken most often at home; ability to speak the official languages; ability to speak other, non-official languages; language spoken most often at work) are only asked in the long form. This latter group of questions is therefore subject to sampling variation as well as to the other types of errors.

A Detailed Discussion of the Canadian Experience, 2001

Statistics Canada regularly produces a series of technical reports pertaining to selected characteristics measured in the preceding censuses of population. The report *Languages* was released in early 2004 (Statistics Canada, 2004). While there is insufficient space here to cover the entire report, it is important enough to include a fairly detailed summary of it. Included are, first of all, the exact question phrasings used in the census. In 2001 only the question on mother tongue was asked in the short form: "What is the language that this person first learned at home in childhood and still understands?" The long form contained five questions:

13. Can this person speak English or French well enough to conduct a conversation?
14. What language(s), other than English or French, can this person speak well enough to conduct a conversation? (Up to three responses were accommodated by the census questionnaire)

15a). What language does this person speak most often at home?

15b). Does this person speak any other languages on a regular basis at home? [Possible responses to this question, other than "No," were: English; French; Other – specify]

16. What is the language that this person first learned in early childhood and still understands?
48a). In this job, what language did this person use most often?
48b). Did this person use any other languages on a regular basis in this job?
[These latter questions were only asked of persons 15 or older].

The processing of the 2001 Canadian census data involved many steps that are not necessary to delineate here. Since many of the responses to the census questions required a respondent to write an answer in a specific place, there was a process by which these write-in responses were coded into computer-readable categories. Most of these responses were converted by an automated process (called "automated coding"), while the remaining responses were coded manually.

Following this coding phase, data were subjected to a process called "edit and imputation," in which missing responses were imputed by a set of complex rules; inconsistent responses were, as much as possible, resolved by a similar set of rules. These two processes dealt with such situations as non-languages such as "Baby" or "Canadian," neither of which are, of course, proper languages, and with pseudo-languages such as "Scandinavian" or "Swiss." I will not dwell on the exact procedures followed in these editing stages. The report gives some overall percentages of the cases affected by these processes. The data on knowledge of official languages show that 96.49% of the responses did not require any editing or imputation. There was a non-response rate of 1.15%; a further 0.67% involved multiple responses; and a further 0.47% had inconsistencies between the response to the question on official languages and those on mother tongue and home language. A further 1.23% was classified as "not applicable."

The 2001 census data on language showed a higher non-response rate for the question on mother tongue than did its two immediate predecessors; for the short form, 2.58% was missing, while the corresponding percentage for the long form was 1.40%. However, there was a wide variation in the non-response rates for the different provinces and territories, from 1.46% for Newfoundland and Labrador (linguistically the most homogeneous province) to a maximum of 4.19% for the Yukon Territory (which has a sizeable population with aboriginal origins).

A second aspect of data quality for the mother tongue question regards multiple responses. The question is phrased very clearly to allow for a single response; moreover, the instructions to respondents stipulate that under certain conditions (e.g., the respondent's having learned more than one language in early childhood) the respondent is to select a specific response (such as "the language you learned first"). However, there was a nontrivial proportion of respondents who either did not want to choose or were unable to choose between two or more languages as their mother tongue. In 2001 that proportion was 4.16% on the short form, 1.29% on the long form. Historically the two forms have had different proportions of multiple responses, with the long form always having the smaller proportion. As with the non-response, there was a wide variation between provinces and territories, with the lowest percentage of multiple responses for Newfoundland and Labrador (0.40% for the short form) and the highest percentage for Ontario (4.88% for the short form). Ontario is the province with the highest proportion of recent immigrants.

The other questions in the 2001 census were subjected to equivalent processes. They, too, had errors of content. The "knowledge of official languages" question had a non-response rate of 1.18% (with the highest territorial rate for the Yukon), while the "knowledge of non-official languages" question had a non-response rate of 2.73% (maximum again for Yukon at 6.91%). Multiple responses to these questions were substantially less common than they were to the question on mother tongue: 0.48% for the former, 0.49% for the latter. It should be recognized that the questions on language knowledge were less inviting to multiple responses than was the one on mother tongue.

Finally, the question on home language was phrased as a two-part question. Part A asked "what language does this person speak **most often** at home?" Part B asked: "Does this person speak any other languages **on a regular basis** at home?" We may be somewhat concerned about the reliability of the "on a regular basis" aspect of this question, but the editing procedures cannot tell us to what extent that clause confused respondents. Overall, the non-response rates for part A were 1.40%; for part B, 2.70%. For both of these parts, the non-response was lowest in Newfoundland and Labrador, and highest in the Yukon Territory. Multiple responses to part A were highest for Ontario (2.10%) and lowest for Newfoundland and Labrador. The same pattern was observed for part B, though at a much lower level (only 0.32% overall).

The 2001 census introduced a new language question that addressed the language used at work. Since this was new, there is no earlier information with which to judge the quality of the data. The question was only asked of respondents 15 years and over, and the structure of this question was similar to the one on

the language spoken most often at home. Part A had a national non-response rate of 3.93% and a multiple-response rate of 2.20%. Part B had a non-response rate of 1.56% and a multiple-response rate of 1.90%. In some contrast to the prior questions, the language at work data by province showed the highest error rates for Nunavut (the territory with an Aboriginal majority), Quebec, and New Brunswick. It is not immediately obvious why these provincial patterns differ so markedly from the other ones. This report also contains some interesting information pertaining to the reliability of the census data on language, though strictly on the "language at work" variable. The Quebec government has been concerned with the language choices made by immigrants and in-migrants to Quebec. In order to inform itself about these issues, several studies have been conducted. The report mentions those by Monnier, on the language choices of immigrant workers, and by Beland, on language use in public settings in Quebec. While for both of these studies we can only make aggregate comparisons, it is satisfying to note that the differences between these studies on the one hand, and the 2001 census on the other, are rather small.

Other census bureaus appear to have devoted considerably less effort to the assessment of data quality (at least with respect to the data on language) than has the Canadian census organization. However, with some digging, a few documents may be unearthed that shed some light on these aspects of validity and reliability. I have a report on the Catalan census data for 1991 (Capellades and Olivares, 1995), a short report on the 2001 South African census, and a working paper from the U.S. census bureau dealing with the reliability of the 1980 census data on language (Kominski, 1989). In all of these cases, overall measures of data quality were satisfactory.

General Comments

In addition to the specific categorization of errors discussed above, we should address the more general issue of information quality. In virtually all of the social sciences, these issues are categorized as reliability and validity. We define the former as the extent to which "a particular technique, applied repeatedly to the same object, yields the same result each time." (Babbie and Benaquisto, 2002). The latter can be defined as "the extent to which an empirical measure adequately reflects the real meaning of the concept under consideration" (Babbie and Benaquisto, 2002). Functional equivalents of these definitions are readily found in undergraduate textbooks in any of the social sciences.

By reliability we refer to the extent to which the item we are measuring is measured accurately. In contrast to validity (see below), reliability is strictly a measurement issue. Language questions in censuses and surveys are reliable to the extent that, for example, subsequent measures yield identical results. We do have some evidence of the reliability of census language questions through post-censual surveys in which questions already asked in the census questionnaire are asked again. Responses are reliable to the degree that the two measures give identical results. An early analysis of the reliability of Canadian census data was conducted by Fellegi (1964). The evidence from the Canadian post-enumeration surveys suggests that such questions, at least in Canada, are quite reliable. Brackstone discussed a possible measure of reliability (estimated standard errors) and reported values obtained for the 1971 census (Brackstone, 1978). As he pointed out, such standard errors only address errors that can be classified as random. Systematic errors (bias) are not estimated by standard errors. The 2001 census of South Africa also measured reliability by means of repeated measures. The reliability for these data is given as a consistency percentage: the percentage of respondents who provided consistent (i.e., identical) responses to the same question at two different times.

By validity, we refer to the degree to which the item we are measuring (the indicator) correctly measures the underlying concept. For example, if we were interested in the fluency with which a respondent speaks a given language, we could just ask a question along the line of "how well do you speak X?," with response categories of: Very well, Well, Poorly, Not at all. (Remember that this was the wording of the U.S. census of 2000 in the question regarding fluency in English for persons who speak a language other than English at home). In contrast to reliability, validity is more than just a measurement issue; it also involves underlying theory. We do have some examples of the establishment of the validity of census language questions. Fishman and Terry (1966), in the study of Puerto Ricans in the United States, asked simple self-report questions on bilingualism and other linguistic features; they followed up these measures with finer-grained questions on the underlying linguistic features. Their findings suggest that, at least for the Puerto Ricans they studied, self-report questions on such issues as bilingualism have relatively high validity. Lamy (1978) also conducted a small study in relation to the Canadian census question pertaining to the respondent's ability to conduct a conversation in English and/or French. His method was quite similar to that of Fishman and Terry: on the one hand, the census question on the respondent's ability to speak the official languages was asked. On the other hand, a variety of fluency measures

were administered to assess the degree to which the self-report question validly measured the underlying concept. Here, too, the conclusion was that the self-report questions had adequate validity. We should keep in mind that both of these studies were conducted in situations in which language was not an extremely sensitive political or cultural topic – it stands to reason that the validity and reliability of such data would be of much lower value in situations where language is an element in the struggle for political or social rights. Therefore, one ought to be wary of the Belgian census data of 1947 (see Levy, 1960 for a detailed discussion of this issue).

Surveys

It is impossible within the restrictions of this article to provide sufficient coverage of the availability of language questions in surveys. Obviously, all surveys contain linguistic information, albeit of the most trivial kind. Surveys, like censuses, are self-report instruments, whether the mode of data collection is aural (structured interviews, in person, or by telephone) or written (questionnaires completed with pencil or pen or internet surveys completed by keystrokes). As such, all of these instruments involve the use of at least one language, both in the stimulus (the asking of questions) and the response. In multilingual countries such as Canada, it is therefore useful for an analyst to know in which language the interview was conducted or the questionnaire written. Given the pitfalls of translation, it would also be useful to know in which language the initial instrument was designed. Unfortunately, in many such instances that basic information is unavailable.

However, there are surveys specifically designed to elicit linguistic information from a sample of respondents. Even to enumerate these goes beyond the scope of this article, so I will limit myself to discussing some surveys for which the primary topic was linguistic in some form. In the context of Euromosaic, many surveys were conducted on linguistic minorities in the member states of the European Union. Since these surveys date to 1994, the data pertain only to a proper subset of linguistic minorities in the 15 members of the European Union at the time.

Regrettably, the data collected in most of these surveys are of extremely limited validity, for several reasons. First, the samples are quite small. The reports I have examined indicate a typical sample size of about 300 respondents. While such small totals do allow for reasonably reliable marginal frequencies, any cross-tabulation (such as language fluency by age group) suffers from the problem that cell frequencies are too small for any statistical inference. If the standard errors reported by Brackstone for the 1971 Canadian census data on language are a reasonable guide, frequencies of 100 would have standard errors of about 20. Statistical inference would in such cases report a 95% confidence interval from about 80 to about 120. Smaller frequencies would have proportionally larger standard errors and would therefore be even less reliable.

However, the larger problem lies in the sample design. As far as I can tell, none of the samples was drawn as a probabilistic sample. As such, no statistical inference is possible, nor is any meaningful comparison between surveys possible. Given the observation that for many regional linguistic minorities in the European Union, there simply are no reliable data, it is quite regrettable that the designers of the data collection, as well as the funding organization, spent a lot of time and financial resources on the collection of data that, as it stands, have limited scientific value. (For more information on the Euromosaic project, as well as reports on separate surveys, see Relevant Websites, below.) Not all of the reports are available in English; several are available in French and/or Catalan.

Courbage (1998) discussed the availability of information on languages in official surveys (that is, those carried out by national statistical agencies, rather than by private or academic survey centers) of about 37 European countries. For many of those (e.g., Iceland, Portugal) such surveys would not be very useful, since they essentially contain only one language community. However, in several of the countries in which two or more language communities have coexisted for some time, apparently no official-survey data on language are available. In the list of 37 countries for which information was supplied, only 9 had conducted national surveys containing questions on language (Courbage, 1998): Portugal, Switzerland, Finland, Russia, Romania, Hungary, Latvia, Estonia, France, Poland, Lithuania, and Luxembourg (note: the last three countries did conduct surveys, but not on a national scale).

There is a fairly extensive collection of social surveys containing language questions for the countries of the European Union: the Eurobarometer surveys which have been carried out since 1973 (in the nine countries then comprising the European Community). Many of these surveys do contain information about language use, mother tongue, or the knowledge of specific languages. While the Eurobarometer surveys are based on probabilistic samples, most of them do not deal with linguistic topics; their utility for the analysis of sociolinguistic topics such as language acquisition, maintenance, or shift is limited.

(For the Eurobarometer surveys website, see Relevant Websites, below.)

Conclusion

In such a brief overview of the prevalence of language questions in censuses and surveys, there is undoubtedly some information about data quality that has not been uncovered. In fact, there is no single source in which we can locate all of the censuses containing at least one language question. The United Nations' demographic yearbooks, which periodically provide summaries of this information, have turned out to be seriously deficient (see de Vries, 2004 for details). Researchers interested in the demographic study of languages are, for now, left to their own devices. The situation is even worse regarding linguistic surveys. Other than the extensive annual surveys cited above, it is virtually impossible to mention all of the available surveys with high linguistic content for any one country, let alone for all countries in the world.

See also: Language Attitudes; Language Planning and Policy: Models; Language Processing: Statistical Methods; Migration and Language Planning; Native Speaker.

Bibliography

Arel D (2002). 'Language questions in censuses: backward- or forward-looking?' In Kertzer D I & Arel D (eds.) *Census and identity: the politics of race, ethnicity, and language in national censuses*. Cambridge: Cambridge University Press. 92–120.

Babbie E R & Benaquisto L (2002). *Fundamentals of social research*. First Canadian edition. Toronto: Thomson Nelson.

Bondarskaya G A (1993). 'Nationality in the population statistics in the U.S.S.R.' In Statistics Canada and U.S. Bureau of the Census *Challenges of measuring an ethnic world: science, politics and reality.* Washington, DC: U.S. Government Printing Office. 333–361.

Brackstone G J (1977). 'The Reliability of 1971 Census Language Data.' In Lamy P (ed.) *Language maintenance and language shift in Canada: new dimensions in the use of census language data*. Ottawa: University of Ottawa Press. 23–37.

Capellades J & Olivares M D (1995). 'Els censos lingüístics a Catalunya: metodologia i tractament.' In Generalitat de Catalunya *Actes del simposi de demolingüística.* 37–71.

Clyne M (1991). *Community languages. the Australian experience*. Melbourne: Cambridge University Press.

Courbage Y (1998). 'Survey of the statistical sources on religion, language(s), national and ethnic group in Europe.' In Haug W *et al.* (eds.) *The demographic characteristics of national minorities in certain European states*. Strasbourg: Council of Europe Publishing. 23–42.

Crawford J (2002). *Making sense of Census 2000.* Tempe: Arizona State University, Language Policy Research Unit. working paper EPSL-0208–105-LPRU.

Del Pinal J & Lapham S J (1993). 'Impact of ethnic data needs in the United States.' In Statistics Canada and U.S. Bureau of the Census *Challenges of measuring an ethnic world: science, politics and reality.* Washington, DC: U.S. Government Printing Office. 447–471.

de Vries J (1985). 'Some methodological aspects of self-report questions on language and ethnicity.' *Journal of Multilingual and Multicultural Development 6,* 347–368.

de Vries J (1990). 'On coming to our census: a layman's guide to demolinguistics.' *Journal of Multilingual and Multicultural Development 11,* 57–76.

de Vries J (2004). 'Language censuses.' In Ammon U *et al.* (eds.) *Soziolinguistik 2. Auflage*. Berlin: Mouton de Gruyter.

European Commission (1996). *Euromosaic: the production and reproduction of the minority language groups of the EU*. Luxembourg: Office for Official Publications of the European Communities.

Fellegi I P (1964). 'Response Variation and its Estimation,' *Journal of the American Statistical Association 59,* 1016–1041.

Finnäs F (1986). *Den finlandssvenska befolkingsutvecklingen 1950–1980.* Helsingfors: Svenska Litteratursällskapet I Finland.

Generalitat de Catalunya. Departament de Cultura. (1995). *Actes del simposi de demolingüística.* Barcelona: Generalitat de Catalunya.

Haug W *et al.* (eds.) (1998). *The demographic characteristics of national minorities in certain European States.* Strasbourg: Council of Europe Publishing.

Jiu J (1999). 'Making sense of Census 2000.' *American language review 3,* 1–2.

Joy R J (1992). *Canada's official languages.* Toronto: University of Toronto Press.

Kelly L G (ed.) (1969). *Description and measurement of bilingualism: an international seminar.* Toronto: University of Toronto Press.

Kertzer D I & Arel D (eds.) (2002). *Census and identity: the politics of race, ethnicity, and language in national censuses.* Cambridge: Cambridge University Press.

Kertzer D I & Arel D (2002). 'Censuses, identity formation, and the struggle for political power.' In Kertzer D I & Arel D (eds.) *Census and identity: the politics of race, ethnicity, and language in national censuses*. Cambridge: Cambridge University Press. 1–43.

Khubchandani L M (1995). 'Linguistic censuses around the world.' In *Generalitat de Catalunya Actes del simposi de demolingüística*. Barcelona: Generalitat de Catalunya. 109–119.

Kloss H (1929). 'Sprachtabellen als Grundlage für Sprachstatistik, Sprachenkarten und für eine allgemeine Soziologie der Sprachgemeinschaften.' *Vierteljahrsschrift für Politik und Geschichte 1,* 103–117.

Kominski R (1989). *How good is "How well"? An examination of the census English-speaking ability question.*

Paper presented at the 1989 Annual Meeting of the American Statistical Association, Aug. 6–11. Washington, DC.
Lamy P (ed.) (1977). *Language maintenance and language shift in Canada: new dimensions in the use of census language data.* Ottawa: University of Ottawa Press.
Levy P M G (1960). *La quérelle du récensement.* Brussels: Insitut Belge de Science Politique.
Lieberson S (1966). 'Language questions in censuses.' *Sociological Inquiry 36,* 262–279.
Lieberson S (1969). 'How can we describe and measure the incidence and distribution of bilingualism?' In Kelly L G (ed.) *Description and measurement of bilingualism: an international seminar.* Toronto: University of Toronto Press. 285–295.
Martel A (1991). *Official language minority education rights in Canada: from instruction to management.* Ottawa: Office of the Commissioner of Official Languages.
Statistics Canada (1999). *Coverage. 1996 census technical reports.* Ottawa: Statistics Canada. Cat. 92–370-XIE.
Statistics Canada (2004). *Languages. 2001 census technical reports.* Ottawa: Statistics Canada. Cat. 92–383-XIE.
Statistics Canada and U.S. Bureau of the Census (1993). *Challenges of measuring an ethnic world: science, politics and reality.* Washington, DC: U.S. Government Printing Office.
Stevens G (1999). 'A century of U.S. censuses and the language characteristics of immigrants.' *Demography 36,* 387–397.

Relevant Websites

http://www.uoc.es/euromosaic/index.html – Euromosaic website.
http://www.gesis.org/en/data_service/eurobarometer – Eurobarometer website.

Language Teaching Traditions: Second Language

D Musumeci, University of Illinois at Urbana-Champaign, Champaign, IL, USA

Introduction

Learning a language in addition to one's first, or native, language has its roots in prehistory, when tribes encountered other tribes whose language differed from their own and a need to communicate arose, perhaps to exchange goods, form alliances, or ask directions. With no written records on which to rely, we can only conjecture as to how that learning happened. However, it would not be unreasonable to assume that it occurred in much the same way as uninstructed (sometimes called 'informal' or 'natural') language acquisition happens today, given similar communicative needs. Instructed (or 'formal') second language learning, however, has a long and varied tradition that may or may not have been based on communicative necessity, depending on the particular historical context in which it occurred Traditions in language teaching reflect a mix of earlier, established techniques combined with innovative influences justified by contemporary ideas in philosophy, religion, and later, psychology, in addition to cultural norms and values. A fascinating aspect of language teaching is that particular themes continued to recur throughout its history, down to the present day.

The following overview of language teaching traditions traces their history from classical Greece and Rome to the 20th century. The perspective is Western European based on the role of Latin and, to a lesser extent, Greek in the curriculum. This is not to suggest that other, non-Western traditions do not contribute to our understanding of historical practice. For example, the oral tradition associated with non-Western educational approaches can provide important insights into second-language teaching and learning. Unfortunately, research on the teaching of second languages within and from the non-Western perspective – as opposed to the teaching of Western languages in the non-Western context – is an area that remains largely unexplored in academic research. Reagan (1996) provides a general overview of non-Western educational traditions, including the teaching of first, but not second, languages.

Early Greek Education

Greeks during the 6th century B.C. held two conflicting views of education in the schools of Athens and in those of Sparta. Although the Athenian model formed the basis for the Western tradition, it is nonetheless interesting to review the Spartan model to see which aspects were shared between the two.

Schools in Sparta

Sparta was a military state established in the 8th century B.C. The Spartan citizen lived for the welfare of the state, of which he was the property, and individuals had no importance apart from the state.

Paper presented at the 1989 Annual Meeting of the American Statistical Association, Aug. 6–11. Washington, DC.
Lamy P (ed.) (1977). *Language maintenance and language shift in Canada: new dimensions in the use of census language data.* Ottawa: University of Ottawa Press.
Levy P M G (1960). *La quérelle du récensement.* Brussels: Insitut Belge de Science Politique.
Lieberson S (1966). 'Language questions in censuses.' *Sociological Inquiry 36,* 262–279.
Lieberson S (1969). 'How can we describe and measure the incidence and distribution of bilingualism?' In Kelly L G (ed.) *Description and measurement of bilingualism: an international seminar.* Toronto: University of Toronto Press. 285–295.
Martel A (1991). *Official language minority education rights in Canada: from instruction to management.* Ottawa: Office of the Commissioner of Official Languages.
Statistics Canada (1999). *Coverage. 1996 census technical reports.* Ottawa: Statistics Canada. Cat. 92–370-XIE.
Statistics Canada (2004). *Languages. 2001 census technical reports.* Ottawa: Statistics Canada. Cat. 92–383-XIE.
Statistics Canada and U.S. Bureau of the Census (1993). *Challenges of measuring an ethnic world: science, politics and reality.* Washington, DC: U.S. Government Printing Office.
Stevens G (1999). 'A century of U.S. censuses and the language characteristics of immigrants.' *Demography 36,* 387–397.

Relevant Websites

http://www.uoc.es/euromosaic/index.html – Euromosaic website.
http://www.gesis.org/en/data_service/eurobarometer – Eurobarometer website.

Language Teaching Traditions: Second Language

D Musumeci, University of Illinois at Urbana-Champaign, Champaign, IL, USA

Introduction

Learning a language in addition to one's first, or native, language has its roots in prehistory, when tribes encountered other tribes whose language differed from their own and a need to communicate arose, perhaps to exchange goods, form alliances, or ask directions. With no written records on which to rely, we can only conjecture as to how that learning happened. However, it would not be unreasonable to assume that it occurred in much the same way as uninstructed (sometimes called 'informal' or 'natural') language acquisition happens today, given similar communicative needs. Instructed (or 'formal') second language learning, however, has a long and varied tradition that may or may not have been based on communicative necessity, depending on the particular historical context in which it occurred Traditions in language teaching reflect a mix of earlier, established techniques combined with innovative influences justified by contemporary ideas in philosophy, religion, and later, psychology, in addition to cultural norms and values. A fascinating aspect of language teaching is that particular themes continued to recur throughout its history, down to the present day.

The following overview of language teaching traditions traces their history from classical Greece and Rome to the 20th century. The perspective is Western European based on the role of Latin and, to a lesser extent, Greek in the curriculum. This is not to suggest that other, non-Western traditions do not contribute to our understanding of historical practice. For example, the oral tradition associated with non-Western educational approaches can provide important insights into second-language teaching and learning. Unfortunately, research on the teaching of second languages within and from the non-Western perspective – as opposed to the teaching of Western languages in the non-Western context – is an area that remains largely unexplored in academic research. Reagan (1996) provides a general overview of non-Western educational traditions, including the teaching of first, but not second, languages.

Early Greek Education

Greeks during the 6th century B.C. held two conflicting views of education in the schools of Athens and in those of Sparta. Although the Athenian model formed the basis for the Western tradition, it is nonetheless interesting to review the Spartan model to see which aspects were shared between the two.

Schools in Sparta

Sparta was a military state established in the 8th century B.C. The Spartan citizen lived for the welfare of the state, of which he was the property, and individuals had no importance apart from the state.

Contact with foreigners was discouraged; in fact, free travel outside the state was prohibited. The aim of education was to develop character, not intellectual capacity, and thus to create obedient, courageous, disciplined citizens who were physically fit and loyal to the state. Girls received no formal education; they learned at home the ideals of the state and of homemaking. They also studied gymnastics to enhance physical fitness, essential for healthy reproduction. Boys became wards of the state at age 6 years, when they left their family homes and moved into military barracks in units of 64 peers. They would not live in a home situation again until marriage at age 30 years. A state official, the *paidonomous*, supervised education, with the aid of assistants who conducted the actual training. At age 18 years, young men became cadets, at which time they began a 2-year period of training in military strategy and tactics, followed by a 10-year obligatory military conscription. On successful completion of service, men were granted full citizenship.

The Spartan curriculum consisted of extreme physical training, sports, and military drill, motivated by competitiveness and harsh discipline. Training in literature was limited to accounts of military heroism. Language teaching, first or second, was not part of the curriculum. In fact, the adjective 'laconic' derives from the geographical name Laconia, of which Sparta was capital, to describe the terseness of speech for which its citizens were famous.

Schools in Athens

In contrast to education in Sparta, the aim of education in wealthy and cultured Athens, victorious in the Persian War (479 B.C.), was to produce a well-balanced individual, intelligent and of strong moral character. Although the state supervised and regulated elementary education, it did not financially support it. Each school was independent and privately operated by a teacher. Education was not compulsory. On the other hand, a 2-year period of military training, *ephebia*, beginning at age 18, was mandatory. Only sons of free citizens were educated; girls received their education at home from their mothers. From ages 7 to 14, boys studied reading, writing, music, and gymnastics, with different teachers for each subject. Reading instruction began with learning the alphabet. Children sang an alphabet song and formed the letters with their bodies while the rest of the class guessed the letters and words.

All teachers were male, and each pupil was accompanied by a *paedagogos*, a male slave who served the mixed functions of nurse, chaperon, and tutor throughout the school day. The elementary school teacher was among the lowest-status occupations in Athens. Itinerant teachers provided instruction beyond the elementary school, offering curricula dependent on their interests and expertise: grammar, composition, rhetoric, literature, music, mathematics, astronomy, or physics. When military training became voluntary, it was replaced by higher education in philosophy, rhetoric, and science, offered through private academies.

Plato and the Academy

A military hero from an aristocratic and wealthy family, at age 20, Plato became a student of Socrates, studying with him for almost 8 years until he witnessed his teacher's trial and conviction in 399 B.C. Disillusioned with Athenian democracy, Plato left Athens, only to return in 387 B.C. and purchase a recreation grove dedicated to the god Academus, wherein he opened a school, the Academy. Plato charged no tuition, relying on the donations of wealthier students to support the enterprise. Both men and women were welcome to study at the Academy. However, only advanced students – those who had already studied geometry – were accepted. The teaching method was lecture-based, with some elements of Socratic dialogue. The curriculum included higher mathematics, astronomy, music, literature, law, history, and philosophy. Plato's epistemology (theory of the nature of knowledge) viewed knowledge as a recalling of ideas that are innate in the soul. He believed that man does not arrive at truth through the senses or by experience. Instead, he must turn inward, looking inside himself. In this way, he can arrive at innate truths through reason. 'Education,' literally 'to draw out of' derives from the idea that learning is the recollection, or remembering, of what is already known and that exists within.

Plato's philosophy, translated into education theory, held that all children should be educated to the limits of their abilities, and that the state, rather than the family, should provide that education. The aim of education was to produce individuals (rulers, warriors, workers, and civil servants) who were oriented to their role in society and whose characters were disciplined to control their animal appetites; that is, to subordinate their senses to reason. Until ephebia at age 18 years, education was devoted to the study of mathematics, literature, poetry, and music. Plato recommended that elementary-level learning be as close to play as possible and that higher levels of learning develop students' critical thinking skills and their ability to use abstract reasoning.

Aristotle and the Lyceum

Aristotle became the most famous of Plato's students in the Academy. At age 41 years, he became tutor to

Alexander, son of King Philip of Macedon, who would later become known as Alexander the Great. At age 50 years, Aristotle returned to Athens and, following in the footsteps of his famous teacher, purchased property and opened a school. The property, dedicated to Apollo Lyceus, provided the name for the school, the Lyceum. The Lyceum became known for its work in natural sciences and was the site of the first zoo and botanical gardens in the Western world. Aristotle's keen observations of nature – honed over the years by examining samples of animal and vegetable life brought back from Alexander's conquests – became the world's chief source of scientific knowledge for the next 1000 years. Aristotle's teaching style consisted of a morning walk through the gardens with his regular students, during which they exchanged and discussed ideas. School became known as 'peripatetic,' or 'walking about.' After eating lunch with his students, Aristotle gave public lectures on politics, literature, and philosophy. The students organized themselves and performed the administrative duties of the Lyceum. All students were expected to engage in historical or scientific research, much of which formed the basis for Aristotle's propositions.

Like Plato, Aristotle believed that man is a rational animal: an animal because he possesses a body with physical needs and appetites, and rational because he has a soul. Unlike his teacher, who held that man is born with preformed ideas, Aristotle proposed that man is born devoid of knowledge, a *tabula rasa* ('blank slate') and that he formulates ideas as a result of contact with material objects. Whereas Plato would have defined learning as 'education,' learning for Aristotle was a matter of instruction ('to put into'): a process of putting knowledge into an empty, but receptive, mind.

The aim of education under Aristotle was to produce a good man; that is, to change a man who is not good by nature to one who controls his animal activities through reason. Both his intellectual and his physical abilities should be developed to their fullest potential. Women were viewed as inferior to men, and their proper functions, as wives and procreators, were fulfilled in the home through training in the domestic arts and gymnastics. Even among men, education was aristocratic; that is, limited to the sons of citizens. The curriculum was not to serve any vocational function, as such activities were the provenance of slaves. Reading, writing, mathematics, natural science, physical education, and humanities (rhetoric, grammar, poetry, politics, and philosophy) formed the curriculum. Because man learns from nature, by habit, and by reason, the teacher's function consisted of organizing the material in a logical manner. Repetitive drill was used to reinforce what was understood by reason, and correct habit formation was essential in the learning process.

The opposing ideas of Plato and Aristotle, in simplified terms of 'education' versus 'instruction,' or of innate knowledge as opposed to knowledge derived from experience, had a profound influence on Western education, including traditions of language teaching. Twenty-first century debates surrounding the extent to which language acquisition is a function of innate human faculties or a result of environmental factors continue to capture the attention of linguists and to influence teaching practice.

Roman Education

In the Roman Republic (508–146 B.C.), education took place at home. Mothers or older relatives tutored young children. Strict obedience was valued. Children were expected to acquire an elementary knowledge of reading and writing. Instruction, whether in literacy or in the trades and professions, was through apprenticeship; that is, through example and imitation. Education was largely vocational, not erudite.

The Roman conquest of Macedonia and Greece (201–146 B.C.) had a powerful influence on Roman society and education. The acquisition of the new territory brought thousands of well-educated Greeks to serve as slaves in Roman households, where they became the teachers of Roman youth. Greek language, culture, and philosophy, including principles of education, spread. The Greek language was so commonly used among educated people that one could address the Roman senate in Greek and be understood. The inclusion of jokes and plays on Greek words in Roman theater provide evidence that even lower classes of Roman society were familiar with the language. By the beginning of the 3rd century B.C., Greek was the language of prestige and culture among educated and upper–social class Romans, existing alongside Latin in a bilingual society.

Considered more practical in orientation than the Greeks, the Romans viewed education as a means to an end: It conferred prestige, but more important, it led to higher status and, thus, to better marriage prospects and more opportunities for advancement.

The system of formal education in Rome in the first century B.C. was divided into four levels. The first, or elementary, level the *Ludus* (meaning 'games' or 'play'), enrolled children from ages 7 to 12. In it, the *magister* ('teacher') taught reading, writing, and arithmetic. Despite its name, the Ludus was infamous for its harsh discipline, and teaching was primarily by rote. Children depended on memory to learn. Elementary schools were open to children of all free

families, boys and girls alike. In this sense, Roman education was more public than the Greek. (Girls' public education, however, ended at this level.) Romans debated the advantages of public versus private school for their children, and upper-class families employed private tutors for their children. The second level of formal schooling was the Grammar School, enrolling students ages 12 to 16. 'Grammar,' from the Greek *grammatike* ('letter'), was the art or technique of writing, not the compendium of rules that the term implies today. The students who attended Grammar School did not typically come out of the Ludus. Rather, they were students who had been privately instructed at home for their elementary education. Both Latin and Greek grammar schools were available, and students could attend one or the other, or both. It is important, from a language learning perspective, to remember that the grammar schools used Latin or Greek as both the content of the curriculum and the medium of instruction, in what we would consider today an immersion-type setting. Young people who attended them would have already been at least functionally competent in the language, having learned to understand, speak, and probably read it at home from a private tutor. The teacher, or 'grammarian,' taught grammar (composition) by means of literature, primarily through lecture. Lecture (literally, 'reading') consisted of the teacher's reading aloud of literary texts and providing comments. The texts provided both the content of the lesson and the form that students were to imitate. Students took notes and memorized lectures. From ages 16–20 years, students attended the School of Rhetoric, where they learned how to use language effectively through the continued study of grammar, argument, and speech (or oratory, literally 'pleading from the mouth'). The chief purpose of the school of rhetoric was to train students to be successful public speakers. The last and highest level of schooling was the University. Two universities were established in the early years of the Roman Empire: one in Athens and the other in Rhodes, both Greek-language institutions. In addition to higher learning, attending the university would have been a study abroad experience. Students who attended the university could be anywhere from 21 to 45 years old. The principal subject was philosophy, but other subjects included law, mathematics, medicine, architecture, and rhetoric. The well-educated Roman was bilingual in Latin and Greek.

Quintilian

One of the most well-known and influential educators in Rome was Quintilian (35–96 A.D.). After training in rhetoric and pursuing a career in law and politics, he was appointed the first state professorship of rhetoric in Rome. Quintilian authored the *Institutio Oratoria* ('Education of the Orator'), a 12-volume series that covered wide-ranging topics in education from preschool to advice for the practicing orator. His advice was that of a distinguished politician and orator regarding the education of boys from upper-class families who were destined to become future leaders. From that perspective, the primary objective of education was to train students to be effective and persuasive public speakers who would then be good public servants. Self-discipline, moral integrity, and social conscience were highly valued attributes. In an age that lacked print media, an esteemed man who could sway public opinion with his oratorical skills held incomparable value for the state. A command of spoken language in conjunction with a background in its literature, history, poetry, music, and philosophy was the mark of a well-educated citizen.

Many of Quintilian's educational principles would be recognized in modern guidelines for practice. Among other things, he advocated that curricular content must be appropriate to the child's ability level (a better predictor for success than age alone); individual differences, both intellectual and physical, among students must be taken into account, with aptitude being an important factor in determining success; a system of rewards to promote learning is more effective than one of punishment; learning cannot be forced, rather, interest, motivation, and persistence are better served through a pleasurable instructional experience; content should be relevant to contemporary situations, and activities should deal with the practical application of that knowledge; and public education is more beneficial than private for the development of social skills.

With regard to language learning in particular, Quintilian suggested that spelling should reflect pronunciation and that games should be used to encourage learning. He advocated the use of wooden blocks in the shape of letters as a way for young children to learn the alphabet and spelling. Above all, he maintained that earlier is better, especially where language is concerned. He recommended that children learn Greek first and Latin second, as the latter would be learned anyway in the context of daily life. Because he believed that learning derives from instruction, he warned that care must be taken to expose children only to excellent and accurate models of language use, both with regard to their caregivers and to the texts they read. Errors, once inscribed on the wax tablet that was the metaphor for the child's mind (Aristotle's *tabula rasa*), were considered difficult, if not impossible, to erase.

Latin Grammars

Aelius Donatus (4th century A.D.) and Caesariensis Priscianus (end of the 5th to the early 6th century A.D.) were Roman grammarians who wrote Latin grammars. Donatus's short grammar, *Ars minor*, was so widely used that any elementary grammar book became known as a 'donat.' It presented the parts of speech in a question and answer format ("What is a noun? A part of speech that signifies by its case a person or thing specifically or generally.") Examples from literature illustrated forms and correct usage. It also contained lists of commonly made errors (alongside the correct forms) and figures of speech. Donatus referred to students' errors as 'barbarisms,' and it is likely that many of them were incorrect spellings based on language they had learned only from dictation, in addition to influences from Vulgar Latin (i.e., the language commonly spoken by the people), which was quite different both from the classical, literary language that students learned in school and from non-Latin dialects. Priscian's grammar, *Institutiones grammaticae*, meant to follow the *Ars minor*, was an 18-book treatise on all aspects of Latin grammar, phonology, morphology, and syntax, filled with quotations from Latin authors. For many students, the examples from Priscian's grammar constituted their only exposure to Latin literature.

The curriculum that Rome had adopted from Greece, namely, grammar, rhetoric, dialectics, mathematics, astronomy, music, and philosophy, remained unchanged for centuries throughout Western Europe. It was a system of education founded on the study of language and designed for students who were already functionally proficient in the language – Greek or Latin – before they began to study it formally.

Education in the Medieval Age

When Germanic tribes invaded, fractured, and conquered the Roman Empire, many of them accepted the culture of Rome and Greece, including Christianity. As a consequence, the Roman church emerged as a dominant influence in Western Europe. The aim of education changed from the development of the educated citizen to the preparation of a man of God, in anticipation of the afterlife. The focus of education turned away from the practical affairs of the world, from sense experience, from physical education, and from external reality. Truth was viewed as absolute: it was not discovered through experimentation but delivered through faith, and it was found only within the Church. Because pupils were inclined toward evil, as a result of original sin, they had to be disciplined and undergo physical punishment to control their evil inclinations.

The spread of Christianity created a conflict between Christian theology and ancient philosophy in education. Liberal thinkers wanted to maintain what was beautiful from the ancient authors, preserving their culture in a Christian form. Theologians were ambivalent in their attitude toward Latin and Latin authors. On the one hand, the Bible and church services were in Latin, so all clergy needed to learn the language. On the other hand, Latin literature, which had served as the model and the method for language teaching, was pagan and a source of possible moral corruption. Because of this, the classical authors were no longer considered appropriate content, especially for young people. Scripture and the writings of the early Church fathers replaced them as models for learning Latin.

Three types of schools predominated during the Middle Ages: the catechetical school, the cathedral school, and the monastic school. Catechetical schools provided an elementary level of education consisting of fundamental doctrines of faith. They were designed for catechumens, that is, possible converts to Christianity, and they provided only what was essential in Latin; namely, the memorization of prayers and scripture passages. The monastic schools, in contrast, trained boys to become monks. They retained the Roman curriculum of the seven liberal arts, divided into the trivium – grammar, rhetoric, and dialectics, or logical argumentation – and the quadrivium – arithmetic, music, geometry, and astronomy. The cathedral schools prepared clergy and sought to provide advanced knowledge of scripture, doctrine, and ritual. This instruction was combined with the study of grammar, rhetoric, literature, geometry, history, and philosophy. Boys who did not intend to become either priests or monks but who wanted a general education attended either the cathedral or monastic schools, where they studied the liberal arts as 'externs.'

Monks were members of a religious order, or community, who vowed dedication to lives of chastity, obedience, poverty, farming, and teaching. They were workers, not contemplatives. The curriculum of the monastic schools, therefore, stressed practical skills. Latin, too, was learned for utilitarian purposes: reading and singing to participate fully in the ritual activities of the church, writing to copy manuscripts (but not necessarily to understand them), rhetoric to be able to teach and preach effectively, and arithmetic to calculate the dates of Easter.

An early church figure, Jerome (340–420 A.D.), having completed a classical education in Rome

and then studied theology, asceticism, Hebrew, and scripture, translated the Bible from Hebrew and Greek into Latin. This was the first Latin bible, known as the Vulgate. It became the official version of scripture for eight centuries. Jerome also founded a monastery and a monastic school in Bethlehem. In a letter dated 403 A.D., he offered advice to a mother on how to raise her infant daughter, much of which can be traced to Quintilian. Jerome stressed the importance of good models and advocated early play with alphabet blocks. He warned against allowing errors to occur and suggested that to ensure accuracy from the very beginning, the mother should guide her daughter's hand as she wrote on a wax tablet or traced letters carved in a board, "so that her efforts confined within these limits may keep to the lines traced out for her and not stray outside of these" (Ulich, 1954: 165). Errors are to be strictly avoided, as "An unused jar long retains the taste and smell of that with which it is first filled" (166). Jerome also advised that the girl learn both Greek and Latin from the very beginning to avoid a non-native-like accent: "For, if the tender lips are not from the first shaped to this, the tongue is spoiled by a foreign accent and its native speech debased by alien elements" (167). Finally, he suggested that the ideal solution would be to send the girl to a monastery.

Another early Church father, Augustine, in his *Confessions*, provided an account of his language-learning experience. He claimed that language developed out of a need to communicate and that he learned his native language by associating sounds and gestures with objects. He then collected these 'signs' and used them to convey his own meanings and desires. Augustine did not know Greek before he went to school and suffered because of it, causing him to hate the language: "The difficulty of learning a strange language did sprinkle as it were with gall all the pleasures of those fabulous narrations. For I understood not a word of it, yet they vehemently pressed me and with most cruel threats and punishments to make me understand it." He compared that experience to learning his first language, without fear or torment, but simply by listening to people talk to him and attempting to convey his own meanings, concluding that "a free curiosity hath more force in children's learning of languages, than a frightful enforcement can have" (147). By his own admission, Augustine loved classical Latin literature, but he criticized the amount of time and effort spent on it and particularly disliked grammar: "men care more to observe the rules of grammar than the laws of God" (149).

The Rise of Universities

Moslem influence during the 11th and 12th centuries, with access to Greek texts in translation, stimulated renewed interest in classical learning. Scholars traveled to Spain and southern Italy to peruse the tremendous libraries that the Arabs had built. In addition, the growth of cities provoked a need for professional training in law and medicine. The university had its informal beginnings where teachers and students came together to learn and debate, much as they had done in Plato's Academy. For their own protection from interference by secular or Church authorities, teachers and students found it necessary to incorporate themselves; hence, the term 'universitas' (guild or corporation). The University of Bologna was the earliest – established in 1088 – and specialized in law; the University of Salerno specialized in medicine, and the University of Paris specialized in the arts.

At the university, one could attain three levels of degrees: the bachelor of arts, the master's, and the doctor's (from the Latin *docere* 'to teach'). The bachelor of arts degree entitled one to continue for a higher degree. The master's and doctor's were earned through the defense of a thesis, demonstrating one's scholarship. The curriculum for the bachelor's degree remained the seven liberal arts. Because of a continued lack of books in the Medieval period, the method of instruction was still by lecture, delivered by either a master's or doctor's candidate who read something he had written and provided commentary. There was no minimum age for attending the university, and it was not unusual for students to be as young as 12 years old.

As in earlier times, a thorough knowledge of Latin was essential for the successful completion of studies at the bachelor's level and a prerequisite for more advanced study. By this point in time, however, Latin was no longer any student's first language. So, to ensure that students would acquire proficiency in Latin, not only for conducting research and writing but also as a means of spoken communication, students were required to use Latin as all times, in and out of class, even in the 'colleges' (student residences): "It has been decreed that the speaking of Latin shall be strictly observed in all the colleges and lodgings, not only by the simple students but also by the bachelors, according to the statutes,... on penalty of a certain fine to be imposed" (Seybolt, 1921: 72). Students were encouraged to report their peers who didn't speak Latin outside of class. In fact, some students were appointed language spies, called 'wolves,' who recorded the names of students who used the vernacular (their native language) instead of Latin. The

students whose names appeared on the lists were summoned and fined.

For conventional and religious purposes during the Medieval Age, Latin remained the language of school even as it became further and further removed from students' linguistic reality outside the classroom. Students did not possess a functional command of Latin; it was a not used by the vast majority of people in the wider community. Latin was a foreign language. Most students did not use it readily and had to be forced to speak it.

Revival of Classical Studies

With the rediscovery of classical authors, there was a renewed enthusiasm for the *studia humanitatis*; that is, the classical Latin education including literature and history. The proponents of the new learning, the humanists, advocated more than just the correction of manuscripts: They proposed a revival of classical learning and culture. They sought to institute Latin as the language of wider communication, much the way that English is used today. The school curriculum, then, continued to be devoted to the liberal arts, but with the insistence that texts of the ancient authors form the content of the curriculum. Students would, once again, learn language in conjunction with content, through exposure to excellent models, and not as a system of abstract rules. The most prominent educators of the day – Vittorino da Feltre, Guarino Guarini, Desiderius Erasmus – exhorted that learning should be pleasant, that harsh discipline was unnecessary and counterproductive, and that errors are artifacts of a developing grammatical system and not a sign of linguistic or moral decay. New attention was paid to the surroundings and comfort of the pupils, as evidenced by Vittorino da Feltre's delightful boarding school, the *Casa Giocosa* (the Playful House), where children learned Latin and Greek in an Italian countryside villa while enjoying fresh air, simply prepared food, and lots of physical exercise and outdoor games. Erasmus, the Dutch humanist, not one to mince words, put it boldly when he wrote in his treatise *On the right method of instruction*,

> I have no patience with the stupidity of the average teacher of grammar who wastes precious years in hammering rules into children's heads. For it is not by learning rules that we acquire the power of speaking a language, but by daily intercourse with those accustomed to express themselves with exactness and refinement, and by the copious reading of the best authors" (Woodward, 1921: 163–164).

Rise of the Vernacular Languages

The humanists' best efforts notwithstanding, they failed in their quest to establish Latin as the universal language. A combination of economic, religious, political, and scientific developments worked against them. As nations formed across the European continent, national languages solidified national identities. Increasing criticism of the abusive power of the Catholic Church stigmatized the use of Latin by association. Although Latin prevailed for a while longer as the language of scholarship and international relationships, it began to lose ground as the vernacular languages grew increasingly powerful. Scientific discoveries began to be published in the vernacular. Galileo published his treatise on planetary movements in Italian, not Latin. The rise of a middle class of merchants and bankers legitimized the vernaculars as media of communication. Parents needed to be convinced of the value of having their children devote so much time and effort to learning Latin. Perhaps the event that had the most significant effect on the language teaching was one of the greatest inventions of all time: the printing press, which allowed for the mass production of books. For the first time, students had easy and relatively inexpensive access to texts. They no longer needed to commit everything to memory or to laboriously copy reams of commentary and lecture notes. Moreover, it wasn't long before the Latin texts were readily available in translation, either interlinear or in side-by-side columns. Such innovation had the obvious effect of eliminating the need for students to struggle through the Latin text to understand its meaning. They could simply read it in their native language.

In addition to his condemnation of the excesses of the Catholic Church, Martin Luther was also a proponent of widespread literacy education. He advocated free elementary education for all children in Germany so that they would be able to read the Bible and thereby attain salvation. Philip Melanchthon functioned as Luther's mouthpiece for educational reform. On the basis of his observations in schools, Melanchthon proposed a three-level system. The plan remained basically humanist, with the innovation that children should first be taught (level 1) in the vernacular before proceeding to the Latin grammar school (level 2) and the conventional course of study, followed by the university (level 3). The rules and regulations that Melanchthon outlined, however, suggest anything but a golden age for learning: those for the university students prohibit a long list of weapons that students were not to bring to their classes, along

with curfews for evenings in the pubs and recommendations to students on how to organize their time effectively.

In England, Sir Thomas Elyot authored the first book on education written and printed in English language, the *Boke named the Governour* (1531). While still advocating the learning of Latin, he argued that English could be used just as well as Latin for scholarly purposes. The Spanish scholar and humanist Juan Luis Vivès (1492–1540) held that language is a living entity, not a static one, and that its use defines its grammar. He proposed that Latin should also be treated as a living language and not simply in imitation of Cicero. Vivès also advocated the use of the vernacular in teaching boys, even though he wrote his treatise in Latin. However, Pietro Bembo (1470–1547), in Italy, not only urged the use of the vernacular but wrote in Italian to praise the Italian language. Bembo was instrumental in establishing as the literary standard the Florentine dialect found in the literary masterpieces by Dante, Petrarch, and Boccaccio.

The Jesuits

The Catholic Church responded to its critics by calling the Council of Trent (1555) to initiate reforms from within. Its program (the Counter Reformation) depended on the education of clergy and laity. Ignatius of Loyola (1491–1556) founded the Society of Jesus, the Jesuits, using a military model: Members would be soldiers who fought for the cause of religion. The Jesuit system of education, although criticized for its elitism, has enjoyed tremendous prestige and esteem since its founding in 1540. The graduate of a Jesuit school was expected to think clearly and logically, express himself eloquently and effectively in speech and in writing, and possess erudition. The *Ratio studiorum* ('Plan of Studies') was the Jesuits' exhaustive description of their educational model. Again, it was a liberal arts curriculum, based on and devoted to the study of Latin. In the lower grammar school, students spent almost 25 hours per week on the study of Latin. Ignatius himself advocated the learning of Latin through literary texts, with Latin as both the medium and the content of instruction. Students were admonished to use Latin at all times. The use of the vernacular was strictly limited, allowed only for the purpose of learning to deliver sermons in it when necessary. Extensive teacher training, rigorous organization, and a carefully prescribed curriculum were hallmarks of the Jesuit system. Despite a lengthy and highly supervised period of training, the least prestigious position was held by the teacher of grammar.

Toward the Modern Era

The esteem with which the Jesuit system was held provided impetus for the Protestants to design a educational system that could compete with it. In opposition to the elitist nature of the Jesuits, the educational reforms proposed by Johannes Comenius included public education for all, regardless of aptitude or intelligence: "a sieve, if you continually pour water through it, grows cleaner and cleaner, although it cannot retain liquid" (Comenius, 1657: 67). A renewed emphasis on observation, experimentation, and reasoning within the scientific paradigm of the day was realized in Comenius's curriculum by a focus on direct experience and learning through the senses. He advocated that pupils study things before words and that teachers organize materials into a natural order, by presenting ideas incrementally, beginning with the known and gradually introducing the unknown, and by recycling material at increasing levels of complexity throughout the curriculum (what he referred to as "the concentric method"). He was also strongly in favor of repetition, explicit error correction, and accuracy from the very beginning: "the first attempt at imitation should be as accurate as possible, that not the smallest deviation form the model be made.... For whatever comes first is, as it were, the foundation of that which follows. If the foundation be firm, a solid edifice can be constructed upon it, but it be weak this is impossible" (Comenius, 1657: 199–200). Although his treatise, *The Great Didactic* (1657) contains many contradictory statements, Comenius's textbooks were his claim to international fame. His major contribution to education is his pioneering use of illustrations as an integral, not merely decorative, element in language textbooks. He, too, advocated elementary instruction in the vernacular school, followed by the Latin school.

Even the most ardent proponents of vernacular education, although advocating education for all, restricted it to the elementary level. Thus they ensured rudimentary vernacular literacy and religious education for the common people. Secondary schools (gymnasia, grammar schools, lycée, academies) were still based on the Latin model and remained the intellectual territory of the elite: boys from well-to-do families who could afford to sent them off to school to be trained in the liberal arts.

Vernacular education at the elementary level predominated in the 17th and 18th centuries, whereas the commonest type of secondary school remained the traditional Latin school. The 'naturalistic' movement in education reflected the major philosophical points of Romanticism and naturalism; namely, an emphasis on emotion as opposed to reason, an intense

interest in nature and intuition, and the belief that the closer man remains to his natural state, the more authentic he is. Rousseau (1712–1778) authored a treatise on education in novel form, *Èmile*, in which he proposed that man is by nature good, but becomes spoiled by the restraints of society and formal education. This position was in direct opposition to the notion that man is born evil and must be saved by God's grace. In line with his philosophical position, Rousseau recommended the elimination of schools altogether.

Êmile, however, influenced the ideas of Johann Basedow (1723–1790) in Germany, who established an experimental laboratory school, the *Philanthropium*. His methodology abolished rote memorization, advocating instead the use of games and a natural, immersion-type approach to teach language, even Latin: "They [the youngsters] played the 'Command' game. You see, it is this way: first, they all stand in a row like soldiers, and Herr Wolke [the teacher] is the officer who commands in Latin, and they must do everything he orders. For instance, when he says *claudite oculos*, they close their eyes tightly: or, *circumspicite*, they peer around in all directions" (Cole, 1965: 429–430) Basedow's curriculum emphasized the importance of the vernacular as the language of instruction and led, ultimately, to the elimination of the Latin grammar schools.

The Lesson of Tradition

When the connection between language and content was severed, Latin became a subject in the curriculum, like science or music, rather than its foundation. Moreover, as it no longer functioned as a means of communication – other than to read ancient texts that were readily available in translation – it served no utilitarian purpose and so could be abandoned for the sake of including more practical subjects in the curriculum. The learning of language for a practical purpose presupposes certain social and economic conditions; for example, international politics, evangelization, commerce, travel, and globalization. In the 20th century, 'living' modern languages replaced Latin in the curriculum, but interestingly, although the language changed, the teaching methodology did not. The legacy that the teaching of Latin left on the early–20th-century curriculum was one of a system of abstract grammatical rules and translation, despite centuries of reformers' advice to the contrary.

In an attempt to create successful classroom conditions for language learning, educators returned, perhaps unknowingly, to the last, most confident era of language teaching – the ideal method – in the tradition of Comenius and the Jesuits. Many methods were introduced: The Silent Way, Suggestopedia, Community Language Learning, and the Berlitz Method. Audiolingualism, a language teaching method popular in the United States in the 1960s, promised to produce competent second-language users through the use of pattern practice, with lots of repetition and absolute accuracy from the beginning. The presentation of language was so rigorously prescribed that learners were not allowed to make mistakes or form bad habits. Touted as a 'scientific' method–based partially on principles of behaviorism in psychology – it prompted schools to make huge investments in language learning laboratories, where students donned headphones, listening and repeating what they heard on tapes. Needless to say, the method did not produce the results it had promised. Unfortunately, it did produce a generation of learners who were convinced that they were incapable of learning language and a cadre of school administrators who were reluctant to invest in future language learning schemes.

The communicative language teaching movement that became popular in the 1970s emerged in opposition to the grammar translation practice of teaching Latin, behaviorist approaches, and any rigidly prescribed method. Proponents of the approach argue that learners acquire language through the interpretation, expression, and negotiation of meaning, rather than through the study of grammatical rules, translation, or mimicry. Others propose immersion education or content-based instruction to maintain the connection between language and content, in a way similar to the bilingual system in early Rome. Others seek to recover the humanist tradition and the centrality of literature in language teaching. Those who take a fully vocational approach suggest curricula designed to teach Language for Special Purposes. Rhetoric and argumentation have resurfaced in Language for Academic Purposes courses.

Consonant with a society that values scientific over humanistic endeavors, some applied linguists look for a scientific orientation to language learning. Such approaches place a renewed emphasis on providing learners with models, in the form of 'input.' Research continues to investigate the extent to which earlier is better, especially with regard to the acquisition of native-like pronunciation. Other linguists take a more philosophical approach and seek to discover universal truths about language and its acquisition, fueling the debate about whether language derives from innate faculties of the human mind or environmental factors.

Language teachers, having never heard of Quintilian, Jerome, or Vittorino da Feltre, readily embrace the role of affect in language learning to explain why

it is important to create an environment that is conducive to learning. They consult language teaching manuals that suggest the use of concrete objects, body movements, and illustrations to convey meaning and to support language development. They include authentic texts in the curriculum to provide models of real language use, although the glossing of text may take the form of hyperlinked text, rather than interlinear translation.

Whether, when, and how to deal with learner's errors remains a concern, as well as how to focus learners' attention on form, without losing sight of meaning. How much and what kind of grammar instruction might enhance learning is still under discussion.

Language teaching tradition reveals that true innovations may be rare, but when confronted with the necessity of implementing instruction that leads to successful language learning, the voice of tradition still echoes in contemporary practice.

See also: Communicative Language Teaching; Interlanguage; Romance Languages.

Bibliography

Boyd W (1966). *The history of Western education.* New York: Barnes and Noble.

Cole L (1965). *A history of education: Socrates to Montessori.* New York: Holt, Rinehart, and Winston.

Comenius J (1657). *The Great Didactic.*

Fitzpatrick E A (ed.) (1933). *St. Ignatius and the Ratio studiorum.* New York: McGraw-Hill.

Garin E (ed.) (1958). *Il pensiero pedagogico dello Umanesimo.* Florence: Giuntine and Sansoni.

Keatinge M W (trans., ed.). (1907, 1910). *The great didactic of John Amos Comenius* (Vols 1 and 2). London: Charles Black.

Kelly L G (1969). *25 Centuries of language teaching.* Rowley, MA: Newbury House.

Kramsch C (1993). *Context and culture in language teaching.* Oxford: Oxford University.

Krashen S (1981). *Principles and practice in second language acquisition.* London: Prentice-Hall.

Lee J & VanPatten W (1995). *Making communicative language teaching happen.* New York: McGraw-Hill.

Lightbown P & Spada N (1999). *How languages are learned. Revised edition.* Oxford: Oxford University.

Musumeci D (1997). *Breaking tradition: An exploration of the historical relationship between theory and practice in second language teaching.* New York: McGraw-Hill.

Reagan T (1996). *Non-western educational traditions.* Mahwah, N.J: Lawrence Erlbaum.

Richards J & Rodgers T (2001). *Approaches and methods in language teaching: A description and analysis.* Cambridge: Cambridge University.

Rousseau J J (1762). *Émile.*

Savignon S (1997). *Communicative competence: Theory and classroom practice.* New York: McGraw-Hill.

Seybolt R F (trans.). (1921). *The Manuale scholarium: An original account of life in a mediaeval university.* Cambridge, MA: Harvard University.

Smith L G & Smith J K (eds.) (1994). *Lives in education: a narrative of people and ideas.* New York: St. Martin's.

Ulich R (ed.) (1954). *Three thousand years of educational wisdom.* Cambridge, MA: Harvard University.

Woodward W H (trans., ed.). (1921). *Vittorino da Feltre and other humanist educators.* Cambridge: Cambridge University.

Language Teaching: History

A P R Howatt, Edinburgh, UK

Introduction

The essence of language is communicative activity occasioned by human contact. While social groups remained small and mobile, speech was the most effective medium through which communication could take place, and the acquisition of spoken language became a 'natural' feature of infant development with similar mental processes available for later reuse if required. Five or six thousand years ago, however, human history took a new turn with the emergence of large and permanent settlements in the great river civilizations of the ancient world including Mesopotamia, Egypt, and northern China. In these and similar instances the imposition of social discipline on large populations entailed the creation of durable sign systems to buttress the ephemeral convenience of speech. Visible symbols were etched, painted, or otherwise inscribed on stone, clay, and other long-lasting substances to create permanent texts from epics to inventories. But writing is not a 'natural' process which can be acquired through contact. It has to be taught and learned.

The writing skills trained in the early scribal schools moved beyond their immediate utilitarian function of recording and preserving factual

it is important to create an environment that is conducive to learning. They consult language teaching manuals that suggest the use of concrete objects, body movements, and illustrations to convey meaning and to support language development. They include authentic texts in the curriculum to provide models of real language use, although the glossing of text may take the form of hyperlinked text, rather than interlinear translation.

Whether, when, and how to deal with learner's errors remains a concern, as well as how to focus learners' attention on form, without losing sight of meaning. How much and what kind of grammar instruction might enhance learning is still under discussion.

Language teaching tradition reveals that true innovations may be rare, but when confronted with the necessity of implementing instruction that leads to successful language learning, the voice of tradition still echoes in contemporary practice.

See also: Communicative Language Teaching; Interlanguage; Romance Languages.

Bibliography

Boyd W (1966). *The history of Western education.* New York: Barnes and Noble.

Cole L (1965). *A history of education: Socrates to Montessori.* New York: Holt, Rinehart, and Winston.

Comenius J (1657). *The Great Didactic.*

Fitzpatrick E A (ed.) (1933). *St. Ignatius and the Ratio studiorum.* New York: McGraw-Hill.

Garin E (ed.) (1958). *Il pensiero pedagogico dello Umanesimo.* Florence: Giuntine and Sansoni.

Keatinge M W (trans., ed.). (1907, 1910). *The great didactic of John Amos Comenius* (Vols 1 and 2). London: Charles Black.

Kelly L G (1969). *25 Centuries of language teaching.* Rowley, MA: Newbury House.

Kramsch C (1993). *Context and culture in language teaching.* Oxford: Oxford University.

Krashen S (1981). *Principles and practice in second language acquisition.* London: Prentice-Hall.

Lee J & VanPatten W (1995). *Making communicative language teaching happen.* New York: McGraw-Hill.

Lightbown P & Spada N (1999). *How languages are learned. Revised edition.* Oxford: Oxford University.

Musumeci D (1997). *Breaking tradition: An exploration of the historical relationship between theory and practice in second language teaching.* New York: McGraw-Hill.

Reagan T (1996). *Non-western educational traditions.* Mahwah, N.J: Lawrence Erlbaum.

Richards J & Rodgers T (2001). *Approaches and methods in language teaching: A description and analysis.* Cambridge: Cambridge University.

Rousseau J J (1762). *Émile.*

Savignon S (1997). *Communicative competence: Theory and classroom practice.* New York: McGraw-Hill.

Seybolt R F (trans.). (1921). *The Manuale scholarium: An original account of life in a mediaeval university.* Cambridge, MA: Harvard University.

Smith L G & Smith J K (eds.) (1994). *Lives in education: a narrative of people and ideas.* New York: St. Martin's.

Ulich R (ed.) (1954). *Three thousand years of educational wisdom.* Cambridge, MA: Harvard University.

Woodward W H (trans., ed.). (1921). *Vittorino da Feltre and other humanist educators.* Cambridge: Cambridge University.

Language Teaching: History

A P R Howatt, Edinburgh, UK

Introduction

The essence of language is communicative activity occasioned by human contact. While social groups remained small and mobile, speech was the most effective medium through which communication could take place, and the acquisition of spoken language became a 'natural' feature of infant development with similar mental processes available for later reuse if required. Five or six thousand years ago, however, human history took a new turn with the emergence of large and permanent settlements in the great river civilizations of the ancient world including Mesopotamia, Egypt, and northern China. In these and similar instances the imposition of social discipline on large populations entailed the creation of durable sign systems to buttress the ephemeral convenience of speech. Visible symbols were etched, painted, or otherwise inscribed on stone, clay, and other long-lasting substances to create permanent texts from epics to inventories. But writing is not a 'natural' process which can be acquired through contact. It has to be taught and learned.

The writing skills trained in the early scribal schools moved beyond their immediate utilitarian function of recording and preserving factual

information in lists and other records. New traditions of text creation grew up giving some of the languages or dialects in which they were written a special 'classical' status, which endured long after the communities which spoke them had died out, making them languages that had to be taught and learned consciously rather than acquired through personal spoken contact. The first of these classical languages was Sumerian, which, like Latin at a later date, was assimilated into neighboring cultures where the need for a prestigious literary language could not be met locally. In other cases, however, languages developed classical varieties of their own by preserving highly valued texts and protecting the language in which they were written from the effects of change wrought by colloquial speech. In the process, they created diglossic situations maintained through educational practices for centuries. Classical Chinese is a well-known example, so also are Sanskrit in India, and Arabic which, since the 7th century C.E., as the classical language of the Koranic texts has coexisted with a growing number of local varieties of colloquial Arabic in the Middle East and North Africa. Such classical varieties can of course be 'spoken' in the sense of consciously 'performed,' and like Latin they may even form the basis for a lingua franca, but they lack native speakers and therefore have to be taught to each new generation, often as a rite of passage into an educated elite.

By contrast, the teaching of 'ordinary' foreign languages for utilitarian purposes was not considered a proper use of expensive resources when they could be 'picked up' through direct contact. In Europe, for instance, Latin was the only language admitted to schools and universities for well over a millennium after it ceased to be a living tongue. It was not until the spread of printing technology after the mid-15th century that any serious demand for a reading knowledge of vernacular languages grew up along with the grammars, dictionaries, and other devices through which they could be taught and learned. Despite this increase in status, the teaching of so-called 'modern' languages did not find easy acceptance in the schools – in part because of opposition by the classicists, but also because it was not easy to find permanent teachers (most were peripatetic native speakers, i.e., temporarily resident foreigners).

Not until the 19th century did modern 'contact technologies' start to emerge which totally transformed the system of world communication and hence the teaching of languages. First was the expansion of steam-driven transport headed by railways that enabled an unprecedented number of speakers of different languages to meet face-to-face, and steamships that clamped the bonds of centralized European control ever more tightly on a growing number of colonial territories around the planet. Next came telecommunications which performed much the same functions without anyone having to move from their offices and culminated in 1901 with Marconi sending his first Morse-coded radio message ('SOS') from England to Newfoundland. More were to follow, the technologies of sound recording (1877 onward) being particularly important in the present context, until the new age of the microcomputer which *inter alia* offers facilities that have to some extent redressed the balance in favor of writing (e-mail, fax, etc.) and raised again the possibility of instant translation.

Utilitarian principles of education have become the norm and with them a reasonably secure place for the teaching of foreign languages. However, the emergence of a 'world community' on the back of mass air transport has increased the demand for a small number of 'superlanguages' with a global status, led by English. The future of many smaller languages has become increasingly precarious, but whether they can be saved through deliberate teaching–learning strategies without a growing native-speaker base is doubtful.

The Ancient World

As noted above, the first instance of organized language instruction in history was the training given in the scribal schools of the ancient world, starting in the cities of Sumer (southern Mesopotamia) and Egypt about 5000 years ago. In neither case was literacy widely distributed among the population at large; it was seen as a specific set of skills shared by qualified craftsmen who had been specially trained for the job. Scribal training was long and demanding (it included the preparation of writing materials for instance), but it was much sought after, and scribes belonged to an envied elite, a few of whom graduated to powerful positions in administrative or priestly castes.

Their geographic proximity suggests that the (slightly) earlier Sumerian work may have influenced the Egyptians. The two systems had some similarities, but technology was not one of them: Sumerian signs were incised with a wedge-shaped ('cuneiform') stylus into tablets of soft clay which were allowed to harden, whereas the Egyptian hieroglyphs and their cursive variants were inscribed on flat surfaces such as papyrus with a brush (later a pen). However, as communication systems they both started by devising pictographic images to represent words whose referents were easy to draw such as *man, woman, tree, sun*, etc. (rather like modern foreign language textbooks). These were later supplemented by more abstract ideographic symbols, but there came a point

when pictures of any kind were impossible (in writing proper names, for instance), and, more generally, there was the crucial problem of inventing new signs. A set of systematic procedures was needed that could generate a readable sign from any spoken word.

At this point, Sumerian and Egyptian diverged; in the former, syllables played an important phonological role whereas the latter was a Semitic language in which words could be written down without specifying the vowels because they were predictable from their consonantal context. Despite its linguistic inappropriacy, Sumerian cuneiform was adapted for use by neighboring Semitic languages about 2500 B.C.E., rendering the system even more difficult to teach and learn. At the same time, the prestige of Sumerian as the original written language turned it into the world's first classical language with a literature, mainly of religious texts, which was studied in the Near East for about a thousand years after the spoken language died out.

The influence of the Egyptian system was even more significant than cuneiform. It was adapted for use by other Semitic languages in the area, and in the case of at least one variant used by the Phoenicians and others living in what is present-day Lebanon the signs came to represent individual consonants rather than sequences, thus creating the first alphabet. The consequent lightening of the learning load was a crucial factor in the success of its subsequent dissemination, including in particular its adoption by the Greeks some time between 1000 and 800 B.C.E. (See British Museum (1990) for articles on above topics.)

At the other end of the learning scale from alphabets are logographic orthographies consisting of large numbers of distinctive signs ('characters'), each representing an individual word. The most famous example is Chinese, the world's longest surviving writing system, which has been in continuous use since the second millennium B.C.E. Despite its drawbacks, it is still intact and regarded as a defining feature of Chinese identity. One reason for this longevity is linguistic: logographic characters suit a key feature of Chinese, namely the dominance of monosyllabic words with no internal structure. The second and more important reason is cultural. In its classical form, the written language came to provide a unifying force in Chinese national life which helped to counteract the disunity of a large number of increasingly diverse spoken dialects. Classical Chinese grew from the decision to maintain the language used in the texts of Confucius (551–479 B.C.E.) as emblematic of a high written culture sustained by a formidable system of competitive written examinations which controlled entry into the rarified circles of imperial administration. The classical language retained its power in this way until the foundation of the Republic in 1911 when moves were made to replace it with a new standard based on the northern (Beijing) dialect known since the 1949 revolution as *putonghuà* ('common language'). A new alphabetic orthography called *pinyin* has had only limited success, however, since its introduction in the 1950s.

Literacy did not make much progress in the ancient world, possibly because the early nonalphabetic systems were too difficult for most people to learn, encouraging the growth of exclusive scribal monopolies that came to exercise considerable political and religious control. As an everyday skill, it was undoubtedly useful for listing, labeling, and the like (Linear B, an early Greek orthography deciphered in the 1950s to much public acclaim, turned out to have been used for little else). There was also a demand from the rich and powerful for monumental inscriptions to commemorate triumphs of various kinds, but as far as most people were concerned, reading had little interest. Communication and fun were well enough supplied by spoken language – there was no need for a replacement.

The Classical World

Greece: c. 800–100 B.C.E.

Some time in the early centuries of the first millennium B.C.E., the Greeks made their second and successful attempt to acquire a system of writing by adapting the consonantal alphabet they saw being used by the Phoenician traders of the Levant. They immediately seized on the advantage of having a small number of signs to learn, but realized that for Greek they would need to represent vowels and they used otherwise superfluous Phoenician signs for this purpose.

Nobody knows exactly when or where this transaction took place, but it is generally agreed that by 800 B.C.E. the Greeks had an alphabet, and the relative ease with which it could be learned helped it to make an impact in different parts of the country. Greece did not suffer from the restrictive practices of an intolerant priestly caste or other forms of scribal protectionism (Goody, 1987). In fact, it was uniquely Greek that an early priority for the new technology should have been to provide a written version of the Homeric epics which were already familiar in their oral storytelling guise. The resulting text (completed about 700 B.C.E.) effectively defined Greek identity for the future and functioned as the indispensable foundation for a groundbreaking approach to education.

As Athens moved toward the democracy for which it was to be famous in the 5th century, a system of education was essential to prepare the future rulers

(i.e., all freeborn males) to participate fully as responsible citizens and to recognize that citizenship brought duties as well as privileges. If 'ordinary' men were to take part in the governance of the city and the dispensation of justice, the people had the right to expect from them a maturity of thought and expression made evident through the effective control of spoken language in the public arena.

The learning of eloquence began in the schools whose prime responsibility was the teaching of grammar, i.e., correct spoken and written Greek, to older children. Later, as young adults, they undertook advanced training in rhetoric and dialectic (the logical structure of argument), building up the three-part curriculum that, as the 'trivium,' was to underpin schooling in medieval Europe. This intellectual continuity also owed much to the writings of Aristotle (specifically on rhetoric and dialectic, for which he devised the syllogistic structure).

For the Greeks, learning grammar meant studying their own language (they rejected foreign languages as mere 'bar-bar' ('barbarian') noises), and it was not an analytic process of taking sentences to pieces. It was rather an investigation of the way language was used by writers of repute: 'grammar is the practical study of the normal usages of poets and prose writers' as Dionysius Thrax (c. 170–190 B.C.E.), a teacher of Greek as a foreign language in Alexandria, said in his *Tekhne grammatike* ('The art of grammar,' c. 100 B.C.E., Kemp, 1986: 346).

The basic pedagogical strategy was mimetic, i.e., the absorption by memorization of substantial amounts of text as the basis for the production of appropriately elevated discourse. The first task was, therefore, as Thrax put it, to develop 'skill in reading (aloud) with due attention to prosodic features' (Kemp, 1986), a priority which was less strange than it would be today (our preference for silent reading derives from medieval monasticism). But there was also a practical reason, namely the correct identification of words and longer syntactic units in texts which were written in *scriptura continua*, i.e., without word separation, (Coulmas, 1999: 455–456). Try, for example, the following English sentence; it should prove helpful to read it aloud using appropriate rhythm and intonation ('prosodic features'):

> 'onereasonformemorizationistobuildupapermanentmentalstoreoftexts' (Coulmas, 1999: 25).

To make things even more difficult, Greek texts were sometimes written in lines running left-to-right and right-to-left alternately (so-called *boustrophedon*: lit. 'ox-turning').

Teaching grammar from texts rather than from rules ('inductively' rather than 'deductively,' to use more recent terms) reflected the lack of analytic grammars, a priority which needed to be revised when Alexander's acquisition of an empire in the Near East created a demand for Greek as a foreign language. Dionysius Thrax's *Tekhne grammatike,* already referred to above, contains an account of the famous eight 'parts of speech,' and it has been called the world's first pedagogical grammar, but there is some doubt about the historical status of this feature of the book (Kemp, 1986; Law and Sluiter, 1995/1998; Robins, 1997). With so many texts lost in whole or in part, all we know for sure is that the direction of change at this time was toward a greater concern for the teaching of linguistic form, and that this trend continued under the Romans for whom Greek was a second language which every educated Roman was expected to learn.

Rome: c. 100 B.C.E.–400 C.E.

By the time of Dionysius Thrax, the Romans were in full control of Greece itself and were moving to take over the entire Hellenistic world in the Near East (an aim consolidated by their victory over Cleopatra at Actium in 31 B.C.E.). In a gesture of tribute unprecedented among conquerors, the Romans adopted Greek culture as their own and consciously modeled themselves on their predecessors. As a result, education in Rome was a bilingual system in which precedence was given to the study of a foreign language, a pattern repeated later with Latin in Europe, but in a more extreme form. There was in consequence a demand for grammars of both languages, and many were produced though few have survived, and even fewer are complete.

In the present context, the most influential surviving document from the Roman period is the 12-volume *Institutio oratoria* ('Education of an orator') by Quintilian (Marcus Fabius Quintilianus, c. 35–100 C.E.), the leading rhetorician of his time. The book deals in detail with the making of an orator from early childhood onward. Concern for good language, Quintilian argued, began in the home and parents should take care to bring children up speaking Latin correctly and well. Learning Greek was the first major priority, though the time devoted to it should not be excessive. Acquiring it from a native speaker at home before starting school gave children a clear advantage, and educated Greek slaves were regularly purchased for this purpose, often with their freedom guaranteed when the family no longer needed their services.

Formal schooling, a privilege open only to boys, typically began about the age of seven. The first task was learning to read and write (Quintilian was insistent on the value of an elegant and well-practiced hand). Practice was also the key to accurate and fluent reading following the 'method of syllables,' i.e., 'sounding out' syllables of increasing length and building them into words, an approach which remained the norm in Europe until quite recent times. Once basic literacy skills had been acquired, the child (now aged about twelve) was ready to leave his so-called *literator* and enroll with a *grammaticus* who taught him Greek and Latin grammar.

The aims of language education in Rome followed the Greek example, namely the development of effective skills in public speaking based on the correct use of language strengthened by a detailed knowledge of the great writers. As with the Greeks, much of the training took the form of the memorization and recitation of substantial amounts of text. The main difference, however, was the need to learn Greek as a foreign language, a task which entailed the formal analysis of sentences, i.e., the eight 'parts of speech,' which are discussed in some detail in Quintilian I.

After leaving school, the young adult could continue his education with a *rhetor* such as Quintilian himself whose principal objective was to train him how to speak effectively in contexts such as the law courts or the Senate house (the Romans did not share the Greek interest in disputation, so dialectic was less popular). The teaching methods were based on the study of models (Cicero, etc.) followed by specially constructed improvisation exercises. All the emphasis was on simple but well-delivered rhetoric and the avoidance of self-indulgent oratorical ornamentation.

Among the few language teaching texts that have survived are portions of *De lingua latina* ('On the Latin language') written around 45 B.C.E. by Marcus Terentius Varro (116–27 B.C.E) and a pedagogical grammar of Greek from the 2nd century C.E. written by Apollonius Dyscolus. It has a syntactic slant which it seems influenced Priscian the great Latin grammarian of the 6th century (see below).

Eventually, just as the empire was about to collapse, there appeared what must be the most famous textbook in language teaching history: the two-volume *Ars grammatica* ('The art of grammar') by Aelius Donatus, a teacher of Latin in Rome about 350 C.E. (Law, 2003: 4.6). The first volume, the elementary *Ars minor* ('Shorter grammar'), became synonymous with 'teaching grammar' for more than a thousand years. It is short and discusses each of the eight 'parts of speech' in turn following a catechistical design ('What is a noun?' 'It is a part of speech which has case . . .' etc.) plus a few rather sparse examples. The more advanced *Ars maior* ('Longer grammar') also covers pronunciation and takes up specific problem issues. The question-and-answer format of the *minor* allowed the teacher to test ('hear') what had been learned in a systematic manner. A typical school classroom contained all the pupils regardless of level, which meant that they had to work on their own until summoned by the teacher who repeated the questions in the book. The consequences of error were usually unpleasant, and the best way of avoiding painful failure was by rote memorization, with or without understanding.

Donatus, it should be said, was the last representative of imperial Latin and by a curious coincidence he became the teacher of St Jerome (c. 342–420 C.E.) whose translation of the Bible (the so-called 'Vulgate') became the first major work in the patristic Latin which replaced it and came to dominate the educated culture of the Middle Ages.

The Medieval World

The adoption of Christianity as the official religion of the Roman Empire in the 4th century C.E. provided Western Europe with an intellectual coherence and a common language that sustained it during the waves of (often violent) change between 400 and 600 C.E. which swept the ancient world into history. Not only did Latin itself survive, albeit in a modified form suited to its new patristic role, but it spawned a set of vernacular dialects influenced by contributions from incomer languages such as Frankish. Only in the province of Britannia (which excluded the Celtic lands) was the imperial language replaced entirely by the speech of the Germanic tribes that over-ran it after the withdrawal of the legions. Although the influence of Latinate Rome returned eventually (596 C.E.), the linguistic 'distance' between Latin and both Celtic and Anglo-Saxon gave teachers from the British Isles (sometimes called 'insular' scholars) a special insight into many of the problems faced by elementary learners.

The history of language and education in the Middle Ages can be divided into three broadly defined phases. The first (c. 400–800 C.E.) was a postimperial period which, despite political instability, saw the establishment of a monastic system that gave the church a place of refuge and nurture for a variety of purposes including schooling. This process was complete by the time of Alcuin and Charlemagne in the late 8th century.

Phase two (800–1200 C.E.), between Charlemagne and the so-called '12th-century renaissance,' was notable for two developments. The first was the transition to a more secular society during which education

left the monasteries and began to put down roots in the towns and cities. The second was the growth of new professional studies, but always with Latin as the medium of instruction, which reinforced its unifying role across the continent. Eventually, specialist centers emerged, the so-called 'universities,' the first of which were founded in Bologna, Paris, and Oxford about 1200.

During the third phase (1200–1500), the universities (particularly Paris) encouraged a revival of interest in the ideas of Aristotle stimulated by new translations of his work into Latin. Their impact on linguistic studies focused on the relationship between language and the outside world, which was expressed through various 'modes' of meaning (*modi significandi*, see Law, 2003: chapter 8). Eventually, however, this excessively cerebral scholasticism provoked a new humanism which *inter alia* stimulated a more extensive use of vernacular languages, most controversially in the translation of the scriptures. It might have been expected that this passion for 'real life' would reduce the importance of Latin or even extinguish it altogether. However, the opposite happened: the medieval world disintegrated, but the Latin language rediscovered its classical roots and became the leading force for the study of (pagan) literature.

Phase 1: From Donatus to Alcuin (c. 400–800 C.E.)

At a time when the single most stable cultural feature of medieval Europe was the Latin language, learning it was both an achievement in itself and the gateway to further advancement (in fact, to be literate (*literatus*) meant to be able to read Latin).

We have already noted the appearance of Donatus's *Ars grammatica* about 350 C.E. The analytic approach it adopted underlined the shift of pedagogical priorities from literature to language which was characteristic of the postimperial approach to the teaching of Latin as a foreign language. A century and a half after Donatus, this trend was amplified in the work of Priscian, a teacher of Latin from North Africa who worked in Constantinople about 500 C.E. (see Law, 2003: 4.10). His students were mainly speakers of Greek, and their learning priorities were reflected in the design of his 18-volume study, *Institutiones grammaticae* ('the teaching of grammar'). Most of the work deals at length with accidence, but it ends with two books devoted to syntax. Initially, it was too detailed to gain widespread popularity, but, after winning the support of scholars during the so-called 'Carolingian renaissance' in the 8th century, it was accepted as a key work for advanced learners (McKitterick, 1994).

Grammar was not just a 'useful subject,' however. It was the foundation of the so-called 'liberal arts' curriculum which, as we have seen, had its roots in classical times but which was now systematized into a two-stage framework of seven subjects starting with the language-based 'trivium' (grammar, rhetoric, and dialectic) followed by the more advanced number-based 'quadrivium' (arithmetic, geometry, astronomy, and music). Apart from grammar and elementary arithmetic, most of these subjects were beyond the scope of the schools, but they eventually formed the basis for the work of the universities (12th century onward). The intellectual groundwork for the approach was laid by St Augustine about 400 C.E. and expanded by scholars such as Boethius (c. 475–524 C.E.), who devised the term 'quadrivium,' and Cassiodorus (c. 490 to c. 580 C.E.), an early leader of the monastic movement whose *Institutiones* (c. 550) set out the curriculum in some detail. A little later came St Isidore of Seville (560 to 636 C.E.) who devoted the first three books of his 20-volume *Etymologiae* to an account of the seven arts.

The culmination of this first phase of medieval language education came with the work of an Englishman, Alcuin of York (c. 737 to 804 C.E.). In 781, he was hired by Charlemagne as a special educational advisor and based in a monastery in Tours where he founded a school. Alcuin's writings covered all three subjects of the trivium (his work on grammar contributed to the renewed interest in Priscian, for instance), but like many in the early church he had a special interest in dialectic. Dialectical reasoning encouraged the exploration of links between logic and theology, for instance, whereas rhetoric was less relevant in an ecclesiastical context that favoured disputation, ritual, and prayer rather than extempore preaching.

Phase 2: Latin and the Vernacular Languages (800–1200 C.E.)

After the time of Charlemagne, there were two parallel linguistic processes at work which ultimately defined the scope of the school curriculum in early modern Europe. The first was the expansion and diversification of the role of Latin as education became part of a wider and more secular community. Schools for children were only one instance; there were also specialized centers for adult students which provided professional studies in subjects such as law, medicine, and administration. From the late 12th century onward, these centers formed loosely structured societies of teachers and students called 'universities' (*universitas societas*) which attracted customers from many different countries. Latin was therefore needed both as a medium of instruction across a broadening curriculum and as a spoken lingua franca in a polyglot community.

The second strand of change was the new maturity of the vernaculars, their defining moment being the appearance of major literary texts, some of which, like the Homeric epics of old, had their origins in oral storytelling. Among the earliest were *Beowulf* (Old English) and the *Chanson de Roland* (Old French), both committed to writing before 1000 C.E., and they were followed by works in other languages, culminating in Dante's *Divina Commedia* (1307–1321). But these were still early days: the vernaculars did not yet offer a serious educational alternative to Latin. There were, however, a few examples of Latin grammars written in the students' mother tongue, including one in Old English by Aelfric, Abbot of Eynsham (c. 955 to c. 1020) who also composed a bilingual glossary for learners, and a famous *Colloquium* (dialogue) with an interlinear translation to help the teacher (pupils did not normally possess personal copies of texts). For the most part the teaching of Latin followed the well-trodden paths of Donatus and Priscian, but there was an interesting pedagogical novelty in the shape of grammars in verse, including the widely used *Doctrinale* of Alexander of Villedieu (c. 1200).

So far as learning the vernaculars themselves was concerned, it was assumed that they would be acquired through spoken contact as and when they were needed, and any teaching would be informal. Eventually, a few self-study manuals (*manières de langage*, Lambley, 1920) started to appear, but most international travellers had to get by in some form of Latin or invent a private lingua franca of their own.

Phase 3: The Later Middle Ages (1200–1500)

Intellectual endeavor in the later Middle Ages took a special interest in exploring the relationship between rationality as represented in, for instance, the newly 'discovered' works of Aristotle and spirituality as manifest in Christian theology. The result was a complex system of categories which drew attention to links between ways of understanding (*modi intelligendi*) and means of expression (*modi significandi*). Important though the *modistae,* as they were called, were in higher educational circles, their work had little direct influence on the teaching of Latin except perhaps at a very advanced stage.

Eventually, however, the rarified atmosphere of modistic theory reinforced the view that the study of language had turned into a formalistic and abstract cul-de-sac. There was therefore an urgent need to replace the 'artificial' patristic Latin of the church with the 'genuine article' derived from the great classical authors, pagan though they may have been, and by the 14th century the search was on for copies of neglected texts, conducted by scholars such as Petrarch (1304–1374) and Boccaccio (1313–1375).

At the same time there was a growing interest in learning classical languages other than Latin, including both Greek and those of the Near East, Arabic and Hebrew. A major step forward was a beginners' book for Greek called *Erotemata* ('Questions') in a question-and-answer format based on Donatus (Law, 2003: 223–227). Later, Western teachers started to visit Constantinople in order to acquire Greek at first hand, and after the fall of the city to Islam in 1453, native speaker refugees fled to the West.

The pedagogical implications of these changes were far-reaching. Grammar remained the foundation discipline in the schools, but the shift away from the medieval concern with dialectic was accompanied by a broader, text-based interest in classical literature which reinstated the importance of rhetoric. Initially, change was slow, partly because of the familiar problems of textual copying, but, with the arrival of printing in the mid-15th century, texts of every kind started to circulate in increasing numbers, facilitating a range of choices which had not been available before.

The Early Modern World

So far as language teaching in schools is concerned, the early modern period brought little structural change until the late 18th century. Even in Protestant countries where the rejection of Latin represented a major change of church practice, it continued to dominate the school curriculum, with the occasional addition of Greek, and modern languages were ignored if not deliberately excluded. As we have seen, there were two major reasons for the survival of the classics, apart from inertia: first, there was the use of Latin in the universities, and second, the influence of Renaissance humanism which encouraged the reading of classical literature. In this connection we should note that the newly created printing industry freed the reading community from the editorial priorities of the scriptoria. By the 1550s, the output of classical literary texts had far outstripped that of formerly dominant religious works (Febvre and Martin, 1984: 264; Eisenstein, 1993).

Printing also facilitated the production of grammars and other works (dialogue manuals, glossaries, etc.) designed to help adult learners to cope with an increasingly polyglot world of 'national' languages. The most impressive was John Palsgrave's grammar of French, *Lesclarcissement de la langue francoyse* (1530), but there were many more modest contributions including James Bellot's *English Schoolmaster* (1580), the first manual expressly designed to teach

English as a second language. The customers were Huguenot refugees in London, some of whom opened schools to teach French to English children (Lambley, 1920; Howatt, 2004).

For most children, however, 'school' meant 'Latin,' and 'Latin' continued to mean 'grammar,' though in England there was at last a replacement for Donatus. Known informally as 'Lily' (after one of the authors) or the 'Royal Grammar' because its adoption was made mandatory by royal proclamation (1540), it was called *A short introduction of grammar* and consisted of an introductory account in English of the parts of speech (by John Colet) followed by a syntax in Latin (by William Lily). For more fortunate students, however, Latin instruction sometimes included the study of literary texts, and the best-known pedagogical account came from Roger Ascham (1515–1568) in his *Schoolmaster*, published posthumously in 1570. Ascham had moved in high royal circles and his book was designed for the education of the courtly elite who wanted to develop a good literary style and elegance of expression. It became famous for a teaching method known as 'double translation' in which the learner translated the Latin text into his mother tongue and, after carefully comparing the two texts, the original Latin one was removed and he had to recreate it from his English translation. An important feature of this technique was the insight it provided into the workings of the mother tongue at a time when the formal study of the vernacular was unknown.

Life for the average early modern grammar school pupil (aged between 8 and 14) was as dreary and painful as it had ever been: nothing but the memorization of grammar rules with no sign of humanism in any sense of the term. '[I had] miserably lost the sweetest spring-time of life ... on scholastic trifles,' complained the great educationist Jan Amos Comenius (1590–1670) in 1636 in his *Great Didactic* (Sadler, 1969: 6). This sentiment was echoed later in the century (1693) by John Locke (1632–1704) who described the child as 'chained to an oar, seven, eight or ten of the best years of his life to get a language or two [which could] be learnt almost in playing' (*Some thoughts concerning education,* 1693, cf. Axtell, 1968: 254).

Both men suggested radical reforms which had some impact among more enlightened teachers. Comenius for instance attempted to construct a graded set of textbooks designed to provide a coherent curriculum suited to the growing child (taking him 'from the known to the unknown'), but the only text to find much success was his famous illustrated book for beginners called *Orbis Pictus* (1658). Locke suggested a 'natural' approach to teaching foreign languages, including Latin, by conversational methods ('talking it into children'), but his proposals were really aimed at well-to-do families who could afford to hire private tutors rather than use the moribund grammar school system. In the second half of the century, advanced educational opinion influenced by Jean-Jacques Rousseau's *Emile* (1762) even called for the abolition of *all* language teaching to pre-adolescent children (including mother-tongue literacy), claiming that it encouraged a form of empty verbalism which caused positive harm by obstructing the child's direct experience of the natural world. These ideas were too extreme for Enlightenment Europe, but they proved attractive to more romantic 19th-century reformers such as Pestalozzi (1746–1827) and Froebel (1782–1852).

Outside the schools printing transformed the fortunes of the vernacular languages in at least two fundamental ways apart from the obvious expansion of the number and range of available texts, (see Eisenstein, 1993). First of all it encouraged the processes of standardization, which were already in progress for sociocultural reasons, making it possible eventually for the vernaculars to be taught alongside (or instead of) Latin in school. Spelling was the obvious first target, but the process spread to all aspects of language. Some of the major European languages, e.g., French and Italian, were taken over by national 'academies' which issued directives on grammatical correctness and stylistic propriety. Others such as English avoided this fate but developed a standardized form through public support for 'authoritative' texts, including Samuel Johnson's *Dictionary of the English language* (1755), Robert Lowth's anonymous *Short introduction to English grammar* (1762), and John Walker's *Critical pronouncing dictionary* (1791).

Second, there was an increased interest in learning foreign languages, particularly for reading purposes. Initially, grammars of minority languages such as English were in Latin (e.g., John Wallis (1653) and Christopher Cooper (1685), both called *Grammatica linguae anglicanae*), but locally produced bilingual works grew in number during the 18th century. Because there were few teachers of modern languages (they all taught Latin), there was a market for 'compendious' self-study manuals which contained everything: a grammar, including guidance on spelling and pronunciation, dialogues and reading texts, glossaries, practical tourist information, and so on. The need for examples of spoken language (dialogues such as 'at the inn,' etc.) stemmed from an expansion of international travel which had become safer by 1750, if not more comfortable, and increasing numbers, mainly of wealthy Britons, set out on the so-called Grand Tour to Italy by way of France. Because

the Tour could take months or even years to complete, learning a modicum of French and/or Italian was essential.

Working through a language manual was possible for well-educated gentlemen who were familiar with grammar from school, but it was too demanding for most children who needed a graded course of instruction. Eventually, in 1783, J. V. Meidinger published *Praktische französische Grammatik* to teach French to speakers of German. It offered a new method which was much better suited to beginners: the grammar was divided into digestible amounts and presented in a step-by-step fashion along with short bilingual lists of new words. In addition, there were exercises consisting of sentences to translate into and out of the foreign language. This design, which eventually became the norm both for school textbooks and adult self-study manuals, is known to us as the 'grammar-translation method,' but this was a label invented by its detractors in the late 19th century. Modern languages were ready at last to enter the world's classrooms.

The 19th Century

As noted in the Introduction, the 19th century was a period of technological innovation with far-reaching consequences for international communications (transport, mail, telegraph, telephone, etc.) and so for the teaching and learning of languages. Adult learners adjusted to the new situation, but it took a long time for the practicalities of real life to have any serious impact on secondary school language classrooms which continued to be dominated by grammar, literature, and 'the exams.'

Within Europe, for instance, mobility on the railways increased the likelihood of face-to-face contact with foreign language communities and hence a growing need to learn how to speak their languages. Outside Europe, there were two main developments. The first was the growth of European colonialism which tightened its stranglehold with the development of fast modern steamships. In the British Empire (outside India which was a special case), there were attempts by the colonial authorities to provide a limited amount of organized schooling on the model of elementary education at home.

English was normally the medium of instruction but attempts to teach it 'as a foreign language' were out of the question. The only model was the grammar-translation method, and there were no grammars or dictionaries. Schooling began with 'Reading' and 'Writing' and not 'English,' which was a fourth-year subject that included 'parts of speech' and 'literature' just as it did in London. Schools were also run by missionary societies, particularly after the end of the infamous slave trade, and many broke new ground by providing orthographies for their pupils' mother tongues in order to prepare materials for the teaching of reading, but they did little else.

Second, there was a steady increase in the importance of the United States, particularly later in the century when it became the destination for large-scale emigration, which created new needs for English, both among those who went to America and those who stayed behind but wanted to visit their families. As a new immigrant from Germany himself, Maximilian Berlitz (1852–1921) was among the first to realize that such students needed completely different teaching methods to help them to speak a foreign language, and he founded his first language school in Providence, Rhode Island, in 1878 to meet this need.

The 19th century was also a time of intellectual exploration, with a new interest in applying the aims and methods of the sciences to the study of human affairs including language and its history. The science of phonetics was an early manifestation of this interest, and its impact on language teaching from c. 1880 onward was considerable.

Despite all these new needs and opportunities in the outside world, formal education took little notice. New moves in elementary education created improved levels of literacy, but there was no role for languages other than the mother tongue. Foreign languages were for the elite only, and Latin, despite losing its utilitarian role in the universities, maintained its position as a high-prestige subject with massive snob appeal. Modern languages by contrast had to fight for recognition, and the only way they could achieve this was by aping the classics and teaching grammar, translation, and worthy literature. It was not until the last two decades of the century that anybody took the spoken language seriously.

The Grammar–Translation Method

As we saw above, the so-called 'grammar–translation method' had been developed by Meidinger with school pupils in mind, and his approach was widely imitated, e.g., by J. G. C. Fick (1763–1821) who wrote the first course in English for speakers of German in 1793, and by Karl Ploetz (1819–1881), the best-known language textbook writer in continental Europe. By the mid-1800s grammar–translation had become the standard secondary-school method for both modern and classical languages, and it was particularly well-suited to the priorities of written examinations. It was also adopted by the growing adult market which developed a new 'series' approach to textbook publishing. An author devised a 'template'

(which he called his 'method') and then adapted it to an unlimited number of different languages. The best-known examples were Franz Ahn (1796–1865) whose 'new, practical, and easy' method first appeared in 1834, and, a year later, the even more famous H. G. Ollendorff (1803–1865) who promised to teach a new language in six months. Their customers were for the most part self-improving adults often with a practical need for a foreign language, but schoolteachers regarded the books as 'too easy,' and they strongly disapproved of making 'keys' available to check the exercises.

When the grammar–translation method started, it lived up to its aim of simplifying language teaching by presenting one or two new grammar rules plus a short vocabulary list in each lesson and following them with translation exercises ('exercises' were a novel development).

Unfortunately, the simplicity of the early courses was not maintained as the century wore on. The impact of exams with their 'unseen' translations packed with tricky linguistic minutiae encouraged teachers to include 'everything' in case it 'turned up,' and the textbooks followed suit. The reputation of the method was seriously damaged as the grammar rules became more complex, the lists of 'exceptions' longer, the vocabulary more esoteric, and the exercises sillier, e.g., 'The pen of my aunt is in the garden of the Dutchman' (parodied over many years).

Inadequacies of this kind prompted many individuals to put forward new ideas for reform (see Howatt, 2004: chapter 13 and Howatt and Smith, 2000). Among the best-known were Claude Marcel's (1793–1876) proposal for a 'reading first' approach in *Language as a means of mental culture* (1853) and an interesting but rather limited suggestion by François Gouin (1831–1896) that language lessons should be organized around a 'series,' i.e., a piece of sequential behavior such as opening the window (*L'art d'enseigner et d'étudier les langues* ('The art of teaching and studying languages') (1880). None of these 'bright ideas' made much headway. What was needed was a new methodology which was well-grounded in theory and sensible in practice, and one appeared in the form of the so-called Reform Movement (see Howatt and Smith, 2002).

The Reform of Language Teaching

The key feature of the Reform Movement, which began in the early 1880s, was the historic decision to give spoken language educational precedence over literature. Signs of change began to appear as phonetics gained greater acceptance in the academic world and younger language teachers saw it as a practical starting-point for pedagogical reform. This early link with linguistics was fostered by leaders of the Movement across Europe. In Germany, for instance, there was Wilhelm Viëtor (1850–1918) whose pseudonymous pamphlet *Der Sprachunterricht muss umkehren!* ('Language teaching must change direction completely!') started the ball rolling in 1882, in Denmark Otto Jespersen (1860–1943) who provided the best account of the classroom implications of the Movement in *How to teach a foreign language* (1904), in France Paul Passy (1859–1940) one of the founders of the International Phonetic Association (a lasting monument to the success of the Movement), and in Britain, Henry Sweet (1845–1912) whose *Practical Study of Languages* (1899) became an applied linguistic classic.

The reformers also insisted on two further classroom principles: the use of sensible connected texts (no more silly sentences), and the 'inductive' teaching of grammar by drawing it out of a close study of the texts. In addition, bilingual words lists were rejected, and translation, particularly into the foreign language, was to be avoided as far as possible. It is important to realize that the Movement appealed to secondary-school teachers, i.e., nonnative speakers of the languages they were teaching, but phonetics provided the help they needed to put speech first.

The translation issue came to dominate a similar but less academic new approach called the 'direct method,' which was popularized, but not invented, by Maximilian Berlitz in his schools in America and later in Europe (see Howatt and Smith, 2000: V). The Berlitz schools were intended mainly for adults with limited practical needs. They set out to develop a conversational approach which was enhanced by the presence of a native speaker teacher who did not know the students' mother tongue, and if he did, he was required to deny it. The 'don't translate' motto became the trademark of the method and virtually all the classwork was oral. If there was a theory behind this, it was a rather simplified version of the view that acquiring a second language was essentially the same as acquiring the first – hence the need for conversation, the absence of grammar and translation, and a delayed start to reading and writing.

The 20th Century

During the 20th century, education around the world was transformed, and language teaching along with it, as schooling in basic literacy became universal in principle if not always in fact. More generally, education ceased to be preoccupied with the grooming of an elite and shifted permanently to a utilitarian philosophy in which the curriculum was judged in terms

of its relevance to the perceived and potential needs of the whole community. The classics were largely replaced in school by major foreign or second languages and, given the structure of the world order that emerged after 1945, it was inevitable that English would become the first choice in many if not all instances. The following account inevitably reflects this bias.

The Growth of English Language Teaching (ELT)

As the language of the British Empire, English had spread around the globe in the previous three centuries, but in colonial schools it was still perceived as a 'quasimother tongue,' and until the 1920s teachers used methods and materials originally designed for elementary schools in the United Kingdom. Provision, never generous, was supplemented by missionary schools of various denominations which devoted their energies to teaching literacy in the local mother tongues. In Europe by contrast, English had taken its place in the secondary schools as a foreign language alongside French and/or German, and it had been at the cutting edge of the 'applied linguistic' reforms initiated by the Reform Movement, particularly in the teaching of pronunciation. It was, however, only in the adult sector that the purely utilitarian objectives of Berlitz and his imitators were followed in direct-method conversation classes.

These two strands of 'reform' and the 'direct method' came together in the work of Harold E. Palmer (1877–1949) who taught English along Berlitz lines in Belgium from 1902 until the German invasion in 1914. He was then forced to return to London where he renewed an earlier contact with Daniel Jones (1881–1967) who had since become the head of the Phonetics Department at London University and was about to publish his famous *English pronouncing dictionary* (1917). The two worked together for nearly seven years (1915–22), during which time Palmer published a series of books, including *The scientific study and teaching of languages* (1917) and *The principles of language-study* (1921), which established a new approach to practical language pedagogy called the Oral Method, combining his classroom experience with the insights of modern phonetics (see Smith, 2003: II).

After 1918, a significant straw in the postwar wind was the decision of the Japanese government to reform English teaching in order to promote greater spoken fluency. They approached Palmer who accepted a position as a special advisor starting in 1922. The following year he was appointed as the Director of an Institute for Research in English Teaching (IRET) (1923–1936), where through both research and materials development, he helped to create a specialized profession which came to be known as 'English language teaching' (ELT) after the founding of a journal of that name in 1946 under the editorship of a close Tokyo colleague, A. S. Hornby (see Howatt, 2004).

A second interwar development was the emergence of 'English as a second language' (later, ESL) to refer to educational contexts (initially colonial) where the language played a significant role in the learner's environment. The phrase was first used in a research project in Bengal (modern Bangladesh) in the early 1920s which investigated the notion that reading was a more useful teaching objective than speaking in countries such as India. The key argument in the project report (1926) by the director Michael West (1888–1973) was that education should be bilingual, i.e., the mother tongue should come first, but an international language such as English had a useful 'secondary' role to play, particularly in science and technology. This view grew in influence and after decolonization in the 1950s and 1960s it became the norm in countries where English was retained alongside the national language(s). 'Second language' was also used in 'English-speaking countries' such as the United States, the United Kingdom, etc., to characterize the role of English among members of minority language communities.

West's report also contained practical suggestions for teaching materials that stressed the importance of vocabulary control, extending a research theme that was also present in the work of Palmer and others in Britain and America. In 1936, the investigation culminated in a so-called 'general service list' of the 2000 most useful words for reading materials in ELT (not published until 1953). There was also an ambitious attempt by C. K. Ogden (1889–1957) to create an 'alternative language' called Basic English (1930) by using only 850 words of ordinary English. It made a useful contribution to wartime teaching, but its long-term appeal was limited (see Howatt, 2004; Smith, 2003: V).

The third and final thread in the development of ELT was the rising demand for English as a foreign language (EFL) in the United Kingdom itself during the 1930s, mainly from refugees from war-threatened Europe. After 1945, there was a pause until the1960s when TEFL began to grow fast.

Out of these influences postwar ELT developed a recognizable approach of its own which stressed the importance of carefully graded texts containing the new grammar patterns, a limited vocabulary of frequent items and an oral method which presented the new items in 'situations' that made the meaning clear without having to translate too much. It was

assumed that, with minor modifications, the same course could be used successfully by all learners. In this it was influenced by ideas in linguistics and psychology (see below), which came to be called 'applied linguistics' in the United States in 1948. When these disciplines changed direction in the 1960s, language teaching followed suit.

Language Teaching and Applied Linguistics

In the Reform Movement we saw the first attempt to apply ideas from the new linguistic and behavioral sciences, a trend that continued to grow as the academic study of these subjects took root, particularly in the United States where a strong theoretical foundation was laid down between the wars. The chief linguistic characteristic was the collection and description of data, particularly of Amerindian languages, while psychology focused on basic learning processes such as habit formation. Together they created a 'structural' tradition that had a direct impact on language teaching during and after World War II.

In 1943, for instance, the United States launched a large project called the Army Specialized Training Program (ASTP, see Angiolillo, 1947) to train spoken foreign language skills. Following the new ideas, they created a drill-based methodology designed to habitualize linguistic structures, or 'speech patterns.' When adapted into mainstream language teaching by Charles C. Fries (1887–1967) at the English Language Institute at Michigan, it became known as the Oral, later Structural, Approach (Fries, 1945). The term 'applied linguistics' was coined in 1948 and with it classroom techniques such as 'pattern practice,' which were well-suited to the technology of the language laboratory, that became popular in the early 1960s. With the new technology 'structural' methods became known as 'audiolingual.'

During the 1960s, applied linguistics shifted its attention from 'microlinguistic' matters such as identifying phonological and grammatical patterns to a broader concern with the way language is acquired and how it is used in social contexts. Language acquisition was a key issue with the generativist school of linguistics associated in particular with the work of Noam Chomsky from 1957 onward and in particular his controversial claim that human beings are innately endowed with the capacity to acquire language, the obvious question in the present context being whether this process can work more than once. By the 1980s, there was a substantial body of theory and research into 'second language acquisition' (SLA), which influenced language teaching by emphasizing the importance of natural language use in, for instance, comprehension (Krashen, 1985) and problem solving (Prabhu, 1987). The latter proved influential in stimulating interest in 'task-based learning' in the 1990s.

Alongside psycholinguistics grew the development of sociolinguistics which aimed to account systematically for language variation, in particular the relationship between the standard (written) language as taught in the schools and the dialectal variability of everyday speech in the outside community. This concern was reflected in a shift in foreign language pedagogy away from formal linguistic patterning to the functional uses of language in everyday social life. Out of this grew the dominant teaching methodology of the late 20th century: 'communicative language teaching.'

Communicative Language Teaching (CLT)

The starting point for CLT is the use of language in communication – who uses it, how, when, etc. – and the way that these variables are reflected in the linguistic choices people make when they talk, write, etc. An influential contribution to the early discussion of this issue was a series of lectures by the philosopher J. L. Austin (1911–1960) published as *How to do things with words* (1962) in which he proposed that utterances should be regarded as 'speech acts' (suggestions, requests, questions, etc.) which controlled the development of social discourse. When these new functional categories were added to more traditional semantic notions (location, duration, etc.) in the 'notional–functional approach,' they provided an extra dimension to language teaching which was generally welcomed (Wilkins, 1976).

The best-known instance of CLT in this framework was an international project called the Threshold Level funded by the Council of Europe (1971 onward). The aim was to define the functional minimum competence that a learner should have, and to allow for cross-linguistic comparisons ('how do you make a request in English/ French/ Spanish . . . ?' etc.). Beyond the 'threshold,' however, learners had specific personal purposes for their learning which needed a rather different pedagogical approach.

Teaching English, and other languages, for specific purposes (ESP) was the second major strand in CLT, (Widdowson, 1978). Many adult students had pressing linguistic needs in their jobs, university studies, businesses, holidays, etc., and the communicative approach offered techniques for providing course materials and teaching programs. These included: authentic texts for reading and listening comprehension, activities to practice language in use rather than linguistic form, role-play and improvisation to provide experience of spoken interaction, and the

detailed study of specific genres and text-types suited to particular professional interests. 'Team teaching' linking subject specialists and language teachers was another common feature of ESP.

Conclusion

The growing dominance of utilitarian objectives in foreign language teaching has been evident since the late 19th century and is unlikely to change in the foreseeable future. Meanwhile, more cultural aims remain the concern of a diminishing specialist minority. However, if computer-based technologies of instantaneous translation are successful, some of the demands of everyday communication could be devolved to gadgetry, which would allow broader educational purposes to play a more central role in the teaching of languages.

See also: Applied Linguistics: Overview and History; Class Language; Dionysius Thrax and Hellenistic Language Scholarship; Foreign Language Teaching Policy; Interlanguage; Language Acquisition Research Methods; Lingua Francas as Second Languages; Mother Tongue Education: Nonstandard Language; Mother Tongue Education: Standard Language; Pedagogical Grammars: Second Language; Rhetoric: History; Speech Acts; Standardization; Vernacular.

Bibliography

Alston R C (ed.) (1967–1972). *English linguistics 1500–1800* (365 vols). Menston: Scolar Press.
Angiolillo P F (1947). *Armed forces' foreign language teaching.* New York: Vanni.
Axtell J L (ed.) (1968). *The educational writings of John Locke.* Cambridge: Cambridge University Press.
British Museum (1990). *Reading the past: ancient writing from cuneiform to the alphabet.* Introduced by J T Hooker. London: British Museum Press.
Coulmas F (1999). *The Blackwell encyclopedia of writing systems.* Oxford: Blackwell.
Eisenstein E L (1993). *The printing revolution in early modern Europe* (Canto edn.). Cambridge: Cambridge University Press.
Febvre L & Martin H-J (1984). *The coming of the book: the impact of printing 1450–1800.* (Tr. New Left Books, 1976). London: Verso.
Fries C C (1945). *Teaching and learning English as a foreign language.* Ann Arbor: University of Michigan Press.
Goody J (1987). *The interface between the written and the oral.* Cambridge: Cambridge University Press.
Howatt A P R (2004). *A history of English language teaching* (2nd edn.). With Widdowson H G. Oxford: Oxford University Press.
Howatt A P R & Smith R C (eds.) (2000). *Foundations of foreign language teaching: nineteenth-century innovators* (6 vols). London: Routledge.
Howatt A P R & Smith R C (eds.) (2002). *Modern language teaching: the Reform Movement* (5 vols). London: Routledge.
Kelly L G (1969). *25 centuries of language teaching, 500 B.C. – 1969.* Rowley, Mass: Newbury House.
Kemp J A (1986). 'The *Tekhne grammatike* of Dionysius Thrax, translated into English.' *Historiographia linguistica XIII 2/3,* 343–363.
Krashen S D (1985). *The input hypothesis.* London: Longman.
Lambley K (1920). *The teaching and cultivation of the French language in England during Tudor and Stuart times.* Manchester: Manchester University Press.
Law V (2003). *The history of linguistics in Europe from Plato to 1600.* Cambridge: Cambridge University Press.
Law V & Sluiter I (eds.) (1995/1998). *Dionysius Thrax and the* '*Techne Grammatike.*' Münster: Nodus.
Lepschy G (ed.) (1994). *The history of linguistics* (4 vols). London: Longman.
McKitterick R (ed.) (1994). *Carolingian culture: emulation and innovation.* Cambridge: Cambridge University Press.
Palmer H E (1921/1964). *The principles of language-study.* London: Harrap. Republished Mackin R (ed.) (1964). Oxford University Press.
Prabhu N S (1987). *Second language pedagogy.* Oxford: Oxford University Press.
Robins R H (1997). *A short history of linguistics* (4th edn.). London: Longman.
Sadler J E (ed.) (1969). *Comenius.* London: Collier-Macmillan.
Smith R C (ed.) (2003). *Teaching English as a foreign language 1912–1936: pioneers of ELT* (5 vols). London: Routledge.
Stern H H (1983). *Fundamental concepts of language teaching.* Oxford: Oxford University Press.
West M (1926). *Bilingualism* (*with special reference to Bengal*). Calcutta: Bureau of Education, India.
Widdowson H G (1978). *Teaching language as communication.* Oxford: Oxford University Press.
Wilkins D A (1976). *Notional syllabuses.* Oxford: Oxford University Press.

Language, Visual Cognition and Motor Action

R A Zwaan and M Kaschak, Florida State University, Tallahassee, FL, USA

Introduction

From the beginning of the cognitive revolution, thinking about cognition has been dominated by the computer metaphor of mind. Cognition is conceptualized in terms of software that stores, retrieves, and manipulates abstract, amodal, arbitrary symbols (e.g., Pylyshyn, 1984; Turing, 1950). In principle, this software can be instantiated in a number of media, such as the human brain or a computer system. The assumption that the 'software' of cognition could be studied independently of the 'hardware' has had a marked impact on the field. Whereas before the cognitive revolution systems of perception and action held a central place in explaining human behavior, after the cognitive revolution these systems became irrelevant to the study of cognition. If cognition could be implemented and studied on a computer, then the fact that humans have sense organs and effectors is unimportant.

This 'disembodied' conceptualization of cognition has recently come under fire in the constituent disciplines of cognitive science, such as artificial intelligence (e.g., Brooks, 1991; Pfeifer and Scheier, 2000), philosophy (Clark, 1997; Newton, 1996; Prinz, 2001), cognitive psychology (e.g., Barsalou, 1999; Glenberg, 1997), cognitive neuroscience (Damasio, 1999; Pulvermüller, 1999), and linguistics (Langacker, 1987; Talmy, 1988). Theoretical arguments and empirical data have been adduced suggesting that cognition is grounded in perception and action. According to this 'embodied' view, perception, action, and cognition are closely intertwined. Here, we consider the implications of this view for language processing.

The embodied view of cognition conceptualizes language comprehension as the mental simulation of the state of affairs that is being described. These simulations are thought to involve the reactivation of patterns of brain activation that were recorded during the comprehender's interaction with the world (Barsalou, 1999; Damasio, 1999). These experiential traces (Zwaan, 2004) are not abstract, amodal, and arbitrary mental representations but, rather, are sensorimotor in kind. There is an increasing amount of evidence for the use of sensorimotor representations in language processing. This evidence is reviewed below.

Perceptual Processes and Representations in Language Comprehension

Researchers in cognitive linguistics have long argued that linguistic meaning is rooted in perceptual processes (e.g., Lakoff, 1987; Langacker, 1987). Consider a sentence such as, 'The cord runs from my laptop to the cable modem.' This sentence conveys a visual scan path from the laptop to the modem. The alternative wording, 'The cord runs to the cable modem from my laptop' sounds odd because it conveys an interrupted scan path, followed by another one (see Langacker, 2005, for a discussion). Consider also the sentence, 'The book is on my desk in my office on the fourth floor of the Longmire building.' Several variants of this sentence can be constructed that are all equally grammatical. One example is, 'The book is in the Longmire building on the fourth floor in my office on my desk.' However, some variants are awkward. 'The book is in the Longmire building on my desk on the fourth floor in my office,' sounds awkward because it does not convey a linear zooming-in or zooming-out process.

Brain-imaging studies have provided evidence that the processing of words engages the cortical and subcortical structures that are active when the words' referents are perceived. When subjects name words referring to actions associated with objects, brain areas just anterior to areas active during motion perception are activated, and when words referring to colors are processed, sites just anterior to areas active during color perception are activated (see Martin and Chao, 2001, for a review). Also, negative emotion words engage the amygdala, which is active when negative emotions are being experienced (Isenberg *et al.*, 2000). This evidence is consistent with the idea that words activate memory traces of perceptual experiences regarding the words' referents.

Behavioral evidence has provided further evidence that language and perception are closely intertwined. Subjects tend to make horizontal eye movements when listening to narratives describing horizontal motion and vertical eye movements when listening to narratives describing vertical motion (Spivey and Geng, 2001). Readers routinely activate the shape and orientation of objects described in sentences (Stanfield and Zwaan, 2001; Zwaan *et al.*, 2002). For example, on reading 'He saw the egg in the carton,' people name a picture of a whole egg more quickly than they do a picture of a broken egg, whereas the reverse is true on reading 'He saw the egg in the frying pan.' There also is evidence that visual primes

affect subsequent language processing (Boroditksy, 2000) and that visual displays presented concurrently with linguistic input may create interference (Fincher-Kiefer, 2000).

To summarize, linguistic analyses, brain-imaging studies, and behavioral experiments all converge on the conclusion that there is a close connection between language and perceptual processes. The argument can be made that meaning arises in part out of the perceptual simulations of the referential situation rather than out of the manipulation of abstract, arbitrary, amodal symbols (Barsalou, 1999; Glenberg, 1997; Pecher and Zwaan, 2005; Zwaan, 2004). The qualification 'in part' is important, because meaning does not just arise out of the language-mediated activation of perceptual representations and processes. Consider the sentence, 'John pounded the nail into the wall.' If the embodied view is correct, understanding this sentence involves not only the visual representation of a horizontal nail (Stanfield and Zwaan, 2001) but also an activation of motor representations and processes involved in pounding nails. In other words, the hypothesis is that meaning arises out of the simulation of actions in addition to the simulation of perceptions.

Motor Representations and Processes in Language Processing

Evidence for the claim that motor representations play a role in language processing can be drawn from three main sources: behavioral studies on adult language comprehenders, neuroscience, and studies of language acquisition.

Klatzky, Pellegrino, McCloskey, and Doherty (1989) describe one of the earliest experiments to show a relationship between motor activity and language comprehension. In their experiments, participants were asked to strike poses, such as raising their arm to eye level and pinching their index finger and thumb together. When holding these poses, the participants responded to sentences such as, 'Aim a dart.' Klatzky *et al.* (1989) report that participants were faster to respond when the action described in the sentence was consonant with the posture they were currently holding. Glenberg and Kaschak (2002) provide further evidence for the role of motor activity in sentence processing. When reading sentences that described action toward the body (e.g., 'Sam gave you a pen') or away from the body (e.g., 'You gave Sam a pen'), participants were asked to decide if the sentences were sensible. Half of the time, a 'Yes' response required making an arm movement toward the body; the other half of the time, a 'Yes' response required making an arm movement away from the body. Participants were faster to respond when the action described in the sentence was consonant with the action required to make a response. Most important, the participants' direction of response interacted with the direction of the action in the sentence not only for concrete actions (e.g., giving someone a pen) but also for actions that are abstract or metaphorical (e.g., the metaphorical transfer of information in, 'John told you a story').

Additional support for the role of motor representations in language processing comes from work in neuroscience. Pulvermüller (1999) reviews a body of research showing that perceptual and motoric representations underlie the processing of words. For example, processing words denoting actions or instruments (e.g., a hammer) appears to engage neural regions known to be involved in the planning and execution of actions (Grafton *et al.*, 1997; Martin *et al.*, 1996).

That perceptual and motor traces are recruited to build mental simulations of the events described in the language is congruent with research in cognitive development (e.g., Piaget, 1954) and language acquisition (e.g., Landau and Gleitman, 1985; Goldberg, 1998). One of Piaget's seminal contributions to the field of cognitive development was the discovery that the abstract thinking (or, formal operations) of adult cognition has its roots in children's understanding of literal, concrete events. For example, children may learn about abstractions such as 'causation' by observing how movements of their body effects change in the external environment (Meltzoff and Moore, 1997). Concepts such as 'equality' may be learned via experiences such as counting the number of marbles placed in two bowls.

Studies of language acquisition (e.g., Bailey, 1997; Landau and Gleitman, 1985) suggest that children's early acquisition of verb meanings (and, their acquisition of subcategorization preferences for verbs) relies on knowledge of the action that the verb describes. The relationship between action and verb meaning likely arises because the verbs are produced and understood in the context of particular actions (e.g., 'give' is likely to be accompanied by an arm movement to transfer an object from one person to another). Once a child understands what it means to 'give' in terms of literal actions (e.g., extending one's arm to give an object to someone), they can begin to understand metaphorical actions (e.g., transfer of information, as in telling someone a story) in terms of those literal actions (see Glenberg and Kaschak, 2002, for a discussion). Thus, the use of perceptual and motoric information in language comprehension that has been detailed

in psychology and cognitive linguistics may be, in part, a vestige of the processes involved in cognitive development and language acquisition.

Limitations of the Current Research, and Future Directions

Although there are theoretical arguments and an increasing amount of evidence in support of an embodied view of language comprehension, there are major hurdles to be taken. We identify three areas in need of further development.

First, as yet there is no coherent theory that specifies how perceptual and motor processes are integrated during perception. A key factor to consider in this regard is perspective (MacWhinney, 2005). For example, when we read, 'John stirred his coffee,' do we simply activate a dynamic visual representation of the coffee and the cup, of the hand stirring it, or perhaps of both? This is a matter of overt and covert 'visual' attention. However, an equally important question is whether we take the perspective of John or that of an observer. In the first case, we also would activate the corresponding motor representations of stirring a fluid in a cup with a spoon. In the observer perspective, we also might activate motor representations – given that viewing actions activates areas in the premotor cortex that overlap with areas active during self-performed actions (Rizzolatti and Arbib, 1998) – but presumably in a less elaborate fashion. We may further ask whether other sensory experiences are triggered by the sentence, such as the smell, taste, and temperature of coffee. An important challenge will be to identify the role of linguistic cues in perspective activation. Another important challenge will be to examine how background knowledge influences perspective taking. Conceivably, comprehenders are more likely to adopt agentive or observer-like perspectives depending on the extent to which they can recruit relevant experiences for each.

A second area in which theoretical development and data are needed is in the treatment of abstract concepts. It has been hypothesized that our knowledge of abstract concepts, such as 'justice' and 'slander,' is grounded in perception and action by way of metaphorical extension (Lakoff, 1987). For example, it has been suggested that applying the adjective 'warm' to an abstract notion such as relationships is grounded in the physical experience of warmth a child feels when she is being held by a parent. However, it is doubtful that this is sufficient. People have many direct experiences with abstract concepts, such as 'justice,' and without these, the metaphor would have not-ing to map onto (Barsalou and Wiemer-Hastings, 2005). Instead, Barsalou and Wiemer-Hastings suggest that abstract and concrete concepts are similar in that both involve situational context, but that they differ in which elements of the situational context are activated.

A third area in which theory development and data are needed concerns the operation of sensorimotor representations in complex linguistic situations. Currently, evidence in support of the embodied position is drawn from studies of the processing of single words and sentences. This leads to a simple and straightforward view regarding how and when sensorimotor representations are used in language comprehension. This simple view is probably inadequate to account for the role of this information in naturally occurring contexts of language use, such as having conversations and arguments, or reading stories. It will be important to examine what happens to sensorimotor representations in language processing tasks that are more representative of the way that language is used in our daily lives.

In conclusion, there is mounting evidence that language comprehension is best understood from an embodied perspective rather than from a disembodied perspective. Sensorimotor representations are clearly active during language comprehension. These findings notwithstanding, several challenges lie ahead for theories of embodied language processing.

See also: Cognitive Linguistics; Developmental Relationship between Language and Cognition; fMRI Studies of Language.

Bibliography

Bailey D R (1997). When push comes to shove: a computational model of the role of motor control in the acquisition of action verbs. Ph.D. diss., University of California, Berkeley.

Barsalou L W (1999). 'Perceptual symbol systems.' *Behavioral and Brain Sciences 22,* 577–660.

Barsalou, L W & Wiemer-Hastings K (2005). 'Situating abstract concepts.' In Pecher & Zwaan (eds.). 129–163.

Boroditsky L (2000). 'Metaphoric structuring: understanding time through spatial metaphors.' *Cognition 75,* 1–28.

Brooks R A (1991). 'Intelligence without representation.' *Artificial Intelligence Journal 47,* 139–159.

Clark A (1997). *Being there: putting brain, body and world together again.* Cambridge, MA: MIT Press.

Damasio A R (1999). *The feeling of what happens: body and emotion in the making of consciousness.* New York: Harcourt Brace & Company.

Fincher-Kiefer R (2001). 'Perceptual components of situation models.' *Memory and Cognition 29,* 336–343.

Glenberg A M (1997). 'What memory is for.' *Behavioral and Brain Sciences 20,* 1–19.

Glenberg A M & Kaschak M P (2002). 'Grounding language in action.' *Psychonomic Bulletin and Review 9,* 558–565.

Goldberg A E (1998). 'Patterns of experience in patterns of language.' In Tomasello M (ed.) *The New Psychology of Language.* Mahwah, NJ: Erlbaum. 203–219.

Grafton S T, Fadiga L, Arbib M A & Rizzolatti G (1997). 'Premotor cortex activation during observation and naming of familiar tools.' *Neuroimage 6,* 231–236.

Isenberg N, Silbersweig D, Engelien A, Emmerich K, Malavade K, Beati B, Leon A C & Stern E (1999). 'Linguistic threat activates the human amygdala.' *Proceedings of the National Academy of Sciences 96,* 10456–10459.

Kintsch W (1988). 'The role of knowledge in discourse comprehension: A construction-integration model.' *Psychological Review 95,* 163–182.

Klatzky R L, Pellegrino J W, McCloskey B P & Doherty S (1989). 'Can you squeeze a tomato? the role of motor representations in semantic sensibility judgements.' *Journal of Memory and Language 28, 56*–77.

Lakoff G (1987). *Women, fire, and dangerous things: what categories reveal about the mind.* Chicago: University of Chicago Press.

Landau B & Gleitman L R (1985). *Language and experience: evidence from the blind child.* Cambridge, MA: Harvard University Press.

Langacker R W (1987). *Foundations of cognitive grammar (vol. 1), Theoretical prerequisites.* Stanford, CA: Stanford University Press.

Langacker R W (2005). 'Dynamicity, fictivity, and scanning: the imaginative basis of logic and linguistic meaning.' In Pecher & Zwaan (eds.). 164–197.

MacWhinney B (2005). The emergence of grammar from perspective. In Pecher & Zwaan (eds.). 198–223.

Martin A & Chao L L (2001). 'Semantic memory and the brain: structure and processes.' *Current Opinion in Neurobiology 11,* 194–201.

Martin A, Wiggs C L, Ungerleider L G & Haxby J V (1996). 'Neural correlates of category-specific knowledge.' *Nature 379,* 649–652.

Meltzoff A N & Moore M K (1997). 'Explaining facial imitation: a theoretical model.' *Early Development and Parenting 6,* 179–192.

Newton N (1996). *Foundations of understanding.* Philadelphia: John Benjamins.

Pecher D & Zwaan R A (eds.) (2005). *Grounding cognition: the role of perception and action in memory, language, and thinking.* Cambridge, UK: Cambridge University Press.

Pfeifer R & Scheier C (2000). *Understanding intelligence.* Cambridge, MA: Cambridge University Press.

Piaget J (1954). *The construction of reality in the child.* New York: Basic Books.

Prinz J (2002). *Furnishing the mind: concepts and their perceptual basis.* Cambridge, MA: MIT Press.

Pulvermüller F (1999). 'Words in the brain's language.' *Behavioral and Brain Sciences 22,* 253–336.

Pylyshyn Z W (1984). *Computation and cognition: towards a foundation for cognitive science.* Cambridge, MA: MIT Press.

Rizzolatti G & Arbib M A (1998). 'Language within our grasp.' *Trends in Neurosciences 21,* 188–194.

Spivey M J & Geng J J (2001). 'Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects.' *Psychological Research 65,* 235–241.

Stanfield R A & Zwaan R A (2001). 'The effect of implied orientation derived from verbal context on picture recognition.' *Psychological Science 12,* 153–156.

Talmy L (1988). 'Force dynamics in language and cognition.' *Cognitive Science 12,* 49–100.

Turing A M (1950). 'Computing machinery and intelligence.' *Mind 49,* 433–460.

Zwaan R A (2004). The immersed experiencer: toward an embodied theory of language comprehension. In Ross B H (ed.) *The psychology of learning and motivation,* vol. 44. New York: Academic Press. 35–62.

Zwaan R A, Stanfield R A & Yaxley R H (2002). 'Do language comprehenders routinely represent the shapes of objects?' *Psychological Science 13,* 168–171.

Language/Dialect Contact

D Britain, Essex University, Colchester, UK

Introduction

The emergence of sociolinguistic approaches to dialectology in the 1960s represented a sharp break with the traditions of the discipline. Previously, dialect studies had largely focused on the collection of the most conservative speech forms in the community – research attention was largely directed toward rural areas, toward older people, toward men, and toward those with socially and geographically restricted life-paths (Chambers and Trudgill, 1998). The studies which followed deliberately shifted attention away from such conservative language users to consider the urban, the young as well as the old, women as well as men, different ethnic groups and people of differing social statuses – the classic case is Labov's (1966) research in New York. They demonstrated, to highlight just a few of the very important contributions of this new approach, the orderly heterogeneity of language variation, and methods enabling us to examine language change as it was actually happening, as well as to spot the locus of change in the

Goldberg A E (1998). 'Patterns of experience in patterns of language.' In Tomasello M (ed.) *The New Psychology of Language*. Mahwah, NJ: Erlbaum. 203–219.
Grafton S T, Fadiga L, Arbib M A & Rizzolatti G (1997). 'Premotor cortex activation during observation and naming of familiar tools.' *Neuroimage 6*, 231–236.
Isenberg N, Silbersweig D, Engelien A, Emmerich K, Malavade K, Beati B, Leon A C & Stern E (1999). 'Linguistic threat activates the human amygdala.' *Proceedings of the National Academy of Sciences 96*, 10456–10459.
Kintsch W (1988). 'The role of knowledge in discourse comprehension: A construction-integration model.' *Psychological Review 95*, 163–182.
Klatzky R L, Pellegrino J W, McCloskey B P & Doherty S (1989). 'Can you squeeze a tomato? the role of motor representations in semantic sensibility judgements.' *Journal of Memory and Language 28*, 56–77.
Lakoff G (1987). *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press.
Landau B & Gleitman L R (1985). *Language and experience: evidence from the blind child*. Cambridge, MA: Harvard University Press.
Langacker R W (1987). *Foundations of cognitive grammar (vol. 1), Theoretical prerequisites*. Stanford, CA: Stanford University Press.
Langacker R W (2005). 'Dynamicity, fictivity, and scanning: the imaginative basis of logic and linguistic meaning.' In Pecher & Zwaan (eds.). 164–197.
MacWhinney B (2005). The emergence of grammar from perspective. In Pecher & Zwaan (eds.). 198–223.
Martin A & Chao L L (2001). 'Semantic memory and the brain: structure and processes.' *Current Opinion in Neurobiology 11*, 194–201.
Martin A, Wiggs C L, Ungerleider L G & Haxby J V (1996). 'Neural correlates of category-specific knowledge.' *Nature 379*, 649–652.
Meltzoff A N & Moore M K (1997). 'Explaining facial imitation: a theoretical model.' *Early Development and Parenting 6*, 179–192.
Newton N (1996). *Foundations of understanding*. Philadelphia: John Benjamins.
Pecher D & Zwaan R A (eds.) (2005). *Grounding cognition: the role of perception and action in memory, language, and thinking*. Cambridge, UK: Cambridge University Press.
Pfeifer R & Scheier C (2000). *Understanding intelligence*. Cambridge, MA: Cambridge University Press.
Piaget J (1954). *The construction of reality in the child*. New York: Basic Books.
Prinz J (2002). *Furnishing the mind: concepts and their perceptual basis*. Cambridge, MA: MIT Press.
Pulvermüller F (1999). 'Words in the brain's language.' *Behavioral and Brain Sciences 22*, 253–336.
Pylyshyn Z W (1984). *Computation and cognition: towards a foundation for cognitive science*. Cambridge, MA: MIT Press.
Rizzolatti G & Arbib M A (1998). 'Language within our grasp.' *Trends in Neurosciences 21*, 188–194.
Spivey M J & Geng J J (2001). 'Oculomotor mechanisms activated by imagery and memory: eye movements to absent objects.' *Psychological Research 65*, 235–241.
Stanfield R A & Zwaan R A (2001). 'The effect of implied orientation derived from verbal context on picture recognition.' *Psychological Science 12*, 153–156.
Talmy L (1988). 'Force dynamics in language and cognition.' *Cognitive Science 12*, 49–100.
Turing A M (1950). 'Computing machinery and intelligence.' *Mind 49*, 433–460.
Zwaan R A (2004). The immersed experiencer: toward an embodied theory of language comprehension. In Ross B H (ed.) *The psychology of learning and motivation*, vol. 44. New York: Academic Press. 35–62.
Zwaan R A, Stanfield R A & Yaxley R H (2002). 'Do language comprehenders routinely represent the shapes of objects?' *Psychological Science 13*, 168–171.

Language/Dialect Contact

D Britain, Essex University, Colchester, UK

Introduction

The emergence of sociolinguistic approaches to dialectology in the 1960s represented a sharp break with the traditions of the discipline. Previously, dialect studies had largely focused on the collection of the most conservative speech forms in the community – research attention was largely directed toward rural areas, toward older people, toward men, and toward those with socially and geographically restricted lifepaths (Chambers and Trudgill, 1998). The studies which followed deliberately shifted attention away from such conservative language users to consider the urban, the young as well as the old, women as well as men, different ethnic groups and people of differing social statuses – the classic case is Labov's (1966) research in New York. They demonstrated, to highlight just a few of the very important contributions of this new approach, the orderly heterogeneity of language variation, and methods enabling us to examine language change as it was actually happening, as well as to spot the locus of change in the

community. These early and groundbreaking studies, however, were based solely on the analysis of natives to those communities, excluding migrants who made up often large proportions of the population, particularly of urban centers. Labov, for example, excluded well over 50% of New York residents from his sample on the grounds of nonnativeness (see Kerswill, 1993, for a discussion of this). This was, however, common practice in social dialectological studies for at least the first 20 years of the discipline.

This nativeness issue is important for two reasons. First, one could perhaps question the validity of dialect descriptions that exclude nonnatives in highly mobile communities such as large metropolitan cities. Second, and perhaps more fundamentally, excluding nonnatives can mask the very origins of some linguistic changes that are underway in the community. Horvath (1985), for instance, who did *not* exclude Italians and Greeks from her study of the English of Australia's largest city, Sydney, found that it was indeed these groups that were leading certain linguistic changes affecting the whole city. Furthermore, Fox (2005), studying the English dialects of adolescents in Tower Hamlets, an 'East End' suburb of London and supposedly the home of its 'Cockney' dialect, found that a number of the linguistic changes underway in that part of London are being led by the now majority Bangladeshi immigrant community and are spreading to the speech of White (especially) male adolescents there. A failure to consider the recently arrived immigrant group and a concentration on 'native' White adolescents alone would have projected almost unexplainable patterns of variation which represent a very clear and quite sudden break from the 'Cockney' dialect that preceded in this part of London. It appears, then, as if urban sociolinguistics was still analyzing rather idealized, 'tidied up' communities, and that this began to be seen as problematic for an all-embracing conception of the way varieties were evolving, given the levels of mobility, contact, and social heterogeneity that were being (or had been) witnessed in many locales around the world.

Dialect Contact and Koineization

Peter Trudgill's (1986) book *Dialects in contact* provided the impetus not just for research on the outcomes of rather radical contact between different dialects (such as colonial migrations) but also for a reconfiguration and reanalysis of dialect change more generally in the light of speaker mobility. Trudgill's basic argument was that if (usually convergent) linguistic accommodation occurs between speakers of different dialects over a sustained period of time (as it would in contexts, for example, of colonialism, urbanization, and migration), the accommodated forms can gradually become permanent features of the speakers' dialects. When children acquire language in a context of widespread accommodation among the adult community (much of which, it has to be said, is not necessarily particularly successful or accurate or complete), they develop an L1 that comprises a collection of accommodated forms and continue the accommodation themselves, passing on the converged norms to their own children, who eventually crystallize a focused dialect in the originally highly diffuse and heterogeneous speech community. Trudgill began to address the following questions: Given the different proportions of dialects which went into the original mix, what sort of new dialect emerges from the contact? What processes work on the mixture of dialects in order to focus that new variety? What determines the outcome? Research on dialect contact over the past 30 years points to four main processes which, together, are labeled *koineization* (see also Siegel, 1985).

'Leveling' is the term used for the eradication of marked (socially or in its typological sense) or minority linguistic forms in the dialect melting pot that follows population movement and has been shown to be the dominant force accounting for the structures of new dialects. A number of studies have convincingly shown that forms which were brought by a majority of speakers almost always survive to form part of the structure of the new dialect. Andrea Sudbury's pioneering study of the English of the Falkland Islands (FIE) (2000), for example, showed dramatic degrees of leveling between the input stage (mostly varieties of English from southwest England and Scotland, but also from other parts of Southern England) and present-day FIE. Gone (or almost gone), for example, are: rhoticity, initial fricative devoicing, periphrastic *do,* the velar fricative /x/, the /w–ʍ/ distinction, Scottish vowel length, double modals, the distinctive southwestern pronoun system and so on (see Sudbury, 2000: 202–206 for more details).

Our ability, of course, to highlight leveling as a dominant koineization process relies quite crucially on us having detailed and explicit knowledge both of the structures of the dialects in the input and of the demographics of their users. Kuo's (2005) study of the development of Taiwanese Mandarin (TM) (Chinese, Mandarin), which has formed over the half century since the flight of Chinese Nationalists from the Mainland after the Chinese Civil War (1945–1949), reminds us of the importance of understanding who brought which dialects to the mix. The previous literature on TM had suggested strongly that this new variety was different from Standard Beijing

Mandarin because of the second language acquisition failures of the non-Mandarin speaking population of Taiwan (mostly speakers of Southern Min (Chinese, Min Nan)) who were forced to learn and speak Mandarin after the Nationalists arrived from the Mainland. They pointed to the fact, for example, that while Standard Beijing Mandarin had four retroflex consonants in its inventory, Southern Min had none, and so, when learning Mandarin, the Southern Min speakers merged the retroflexes with corresponding nonretroflex consonants, and diffused the nonretroflex consonants to the children of original mainlanders. Through a very careful analysis both of the structure of Chinese dialects in the middle of the 20th century, and census information on the regional origins of the migrant mainlander population, Kuo demonstrated that while retroflexes were common in central Beijing, they were rarely found elsewhere in China, and were in a tiny minority of the migrant population to Taiwan. She argued, therefore, that the lack of retroflexes in TM was a simple result of them not having been brought to Taiwan in sufficient numbers in the first place, and that the few that were brought were swiftly leveled away (see also Britain, 2001).

‘Simplification’ refers to the process that leads to the new dialect being more regular, having fewer categories (such as gender, case, honorifics) and, for example, fewer person/number inflections than the dialects in the original mix. It is neatly exemplified by the many new varieties of Hindi/Bhojpuri which developed as a result of the migration of indentured laborers from the Indian subcontinent to work on colonial plantations in, for example, Fiji, Mauritius, South Africa, Trinidad, and Guyana (e.g., Barz and Siegel, 1988). Gambhir (1981: 259), for example, compared the number-, gender- and person-sensitive present tense system in Indian varieties of Bhojpuri (with over two dozen morphemes) with that of Guyanese Bhojpuri which has just two.

‘Interdialect’ forms are variants which did not exist in the input varieties, but have developed as a result of the acquisition of imperfectly accommodated forms. Usually they are forms intermediate between two or more of the input variants. In the Fens of Eastern England – koineized as a result of in-migration of southeastern and northwestern dialect speakers following marshland reclamation from the 17th century onward – contact between southern [ʌ] and northern [ʊ] variants of the vowel in words such as ‘cup’ and ‘bun’ has resulted in the gradual emergence of an interdialectal, phonetically intermediate [ɤ] form (see Britain, 2002).

Following ‘reallocation’ more than one variant survives from the dialect mix, but each survivor is reassigned to serve new functions. Khamkhong (2004), investigating new dialect formation in Ban Khlong Sathon (BKS), a small settlement in northeastern Thailand that was established in the 1970s and attracted migrants from across the country, found that reallocation accounted for the distribution of the Thai negators *bor* and *mai*. *Mai* is the negator used in Central (and Standard) Thai (spoken to the west of BKS), and *bor* the form used in the Isan (Thai, Northeastern) dialect of Thai spoken to the north and east of BKS, and these forms were preserved by the original migrants. Khamkhong’s analysis of the grandchildren of the original migrants, however, found both forms being used, but each reallocated according to topic, with *bor* being used most and when discussing informal everyday topics, and *mai* reserved for discussions about more formal topics, such as lessons at school, health issues, and the local council.

Three Contexts of Contact

For the remainder of this article I discuss three major contexts in which linguistic varieties come into contact – ‘new dialect formation,’ the consequence of the mixture of distinct dialects in new speech communities, such as in colonial contexts, or as a result of labor movements, New Town formation, land reclamation, or urbanization; ‘second dialect acquisition,’ the ability (or otherwise) of migrants to acquire a second dialect of their own language; and the ‘geographical diffusion of linguistic innovations.’ The literature to date has tended to treat these as three distinct areas of enquiry, but I will attempt here to demonstrate that, in fact, the outcomes of each show remarkable similarities because all three have dialect contact at their core.

New Dialect Formation

It is the development of new dialects following large scale population movements that has provoked most research in the dialect contact literature (see, for example, Gambhir, 1981; Trudgill, 1986, 2004; Barz and Siegel, 1988; Britain, 1997; Sudbury, 2000; Kerswill, 2002; Kuo, 2005) and from which the assessment of the likely outcomes of contact – leveling, simplification, interdialect formation, and reallocation – has been derived. More recently, researchers have been looking in more detail at the likely order in which these expected outcomes develop following mass migrations.

Perhaps the most detailed investigation of new dialect formation has come from research on the origins of New Zealand English (Trudgill, 2004). At the disposal of the researchers was a large set of recordings

made in the 1940s of mostly elderly New Zealanders who represented the first New Zealand-born generation of Anglophones following the mass migrations from Britain and Ireland in the mid–19th century. Trudgill proposed three stages by which a mixture of dialects, over time, crystallizes into a more focused, less diffuse variety.

Stage 1: Rudimentary Leveling This stage occurs, he claimed, among adults at the very earliest points of contact between the speakers in the mix – in the case of New Zealand English at the various "assembly points in the British Isles, on the long boat journey out, and early on in the new location" (2004: 89). An example of such leveling is the absence in the New Zealand recordings of the merger of /v/ and /w/ into /w/ (e.g., [wɪlɪdʒ] for 'village') that was a receding characteristic of some dialects of the south of England in the mid–19th century.

Stage 2: Extreme Variability Children born shortly after arrival in a new speech community (such as Anglophone New Zealand in the mid to late 19th century) are faced with the task of focusing an L1 without a clear target. Trudgill claimed that in such circumstances children draw more on their parents' and other adult varieties than is usually the case, and the consequence of this, in a dialect mixture situation, is that they often use new and unusual combinations of features that derive from a wide range of dialects, and they often differ from each other quite radically too (see Trudgill, 2004: 103–105, for some examples of individual combinations). He also provided evidence of ongoing leveling, this time of forms which survived the journey, but not in sufficient numbers to be transmitted to the next generation.

Stage 3: Focusing through Deterministic Leveling This is the final stage, the stage at which a new, stable, relatively uniform variety emerges. It is at this point, with many of the minor 'competitors' already leveled away at an earlier stage, that certain forms finally win out. Trudgill claimed that, in almost every case, majority or unmarked forms win, and, perhaps controversially, that prestige, stigma and identity play no role in that determination. It is also at this point that reallocation may take place.

Second Dialect Acquisition

A more recent concern in dialect contact studies has been the factors that shape the ability of migrants to acquire a second dialect of their own language (or languages, as we shall see below). In contexts such as these, a speaker's dialect enters into contact with those of their newly adopted community. A number of common threads have emerged from the small amount of research on this topic to date. First, as might be expected from the research on second language acquisition, age of arrival in the community is crucial. Both Chambers (1992), investigating the acquisition of the English of Oxfordshire in England by Canadian children, and Watts (2000), looking at the acquisition of Cheshire English by American children, found that, on the whole, the younger the child on arrival in the new community, the more successful the acquisition of the second dialect. By contrast, Foreman (2003: 264), examining the acquisition of Australian English by North American adults, found that "the majority of the subjects did not acquire any phonetic or phonological aspects of AusE to an auditorily noticeable extent." While it is true that some studies have found some adjustments by adults in such settings, overall the finding of an inverse relationship between age of arrival and ease of acquisition is quite robust.

Second, it has been found that migrants acquire the new lexicon more quickly than the new phonology. Chambers (1992: 680) found, for example, that the group of Canadian migrant children he studied acquired over 50% of the examined lexical forms (such as using 'trousers' instead of 'pants,' 'nappy' instead of 'diaper,' etc.), but only a quarter of the phonological forms studied in their first two years in the nonnative speech community. Further evidence for this pattern comes from work investigating the acquisition of a second dialect of a second language. Laurie Chou (1997) investigated the extent to which Taiwanese were able to shift from the American English they had been taught in Taiwan to British English when working in England. She found that after three years in England her 30 subjects had acquired, on average, 78% of the 'British' forms from the British/American lexical pairs she investigated, but had /r/ deletion (in words such as 'car,' 'port' etc.) in 64% of tokens (compared with 41% among a control group of Taiwanese in Taiwan – suggesting a deletion rate of just 23% after three years).

Third, Chambers provided convincing evidence that 'simple' rules (i.e., dependent on automatic exceptionless distinctions) are more quickly and successfully adopted by second dialect acquirers than 'complex' ones (i.e. distinctions with exceptions, variant forms, and/or abstract linguistic conditioning, phonological splits, new phonemes, etc.) (Chambers, 1992: 682). He compared the acquisition of a voiceless British English intersonorant /t/ (a simple, exceptionless change from a Canadian English voiced [d]) with that of the COT–CAUGHT split – a lexically-diffused, phonologically complex characteristic of southern British English. While one 9-year-old

Canadian child in Oxfordshire acquired both features over 90% of the time, older children acquired the former, simple rule on average in around 45% of possible places, but the latter complex rule on less than one occasion in ten. In an investigation of the acquisition of three such 'complex' rules of Kuwaiti Arabic (Arabic, Gulf Spoken) by Egyptians, Al-Dashti found that female Egyptians acquired the local Kuwaiti [j] form of the variable (dʒ) and the [g] form of (q) but were unsuccessful in acquiring the [ð] form of (ð) (Al-Dashti, 1998), suggesting that some 'complex' rules are harder to acquire than others. Kerswill (1996: 200) proposed a cline of acquisition difficulty from vocabulary at the 'easy to acquire' end to lexically unpredictable phonological rules at the 'difficult to acquire' end as a way of classifying linguistic constraints on acquisition in a more fine-grained way.

Another, perhaps obviously significant, factor in acquisition success has been length of exposure to the target dialect. Chambers (1992), Al-Dashti (1998), Watts (2000), and Foreman (2003), for example, all showed that longer exposure leads to more successful acquisition. Further research is required on other social factors that may influence the acquisition process. Al-Dashti (1998) found significant gender differences in acquisition, with, in both cases, women proving to be better acquirers than men. Foreman's (2003) study of Americans in Australia emphasized the role of social identity in acquisition rates, claiming that 'feeling Australian' and engaging in mostly Australian and few American social network ties contributed to acquisition success. And Phili (2003) demonstrated that, if there are strong motivations to retain affiliation with their original communities, speakers appear to be able to resist shifting at all to new linguistic norms. She investigated speakers from Famagusta in Cyprus who were forced to leave their homes following the Turkish invasion of the north of the island, and who subsequently settled well away from their town of origin in Limassol. She found few differences between the speech of the migrants, thirty years after leaving Famagusta, and that of people still living in villages within the Famagustan dialect area (but on the Greek side of the divide).

Studies both of adults and children acquiring second dialects have provided evidence of interdialect formation. Watts (2000) showed that a majority of the American children, faced with the Cheshire English [ʊ] realization of the vowel in words such as 'cup' and 'mug' produced a vowel that was phonetically halfway between it and an American English [ʌ].

Evidence from the rather small second dialect acquisition literature also suggests, therefore, that contact processes are at work – 'simple' as opposed to 'complex' forms survive, suggesting leveling and simplification, with adult accommodation being less successful and complete than child accommodation, and occasionally instances of interdialect forms emerging.

The Geographical Diffusion of Linguistic Innovations

The spread of new linguistic variants from one place to another appears to operate in one of three different ways.

Wave-Model Diffusion Perhaps the most iconic model of the geographical spread of innovations and the simplest, since it depends solely on the friction of distance, is the 'wave' or contagion diffusion model whereby innovations, over time, radiate out from a central point, reaching nearby locations before those more distant. Trudgill, for example, in a discussion of the diffusion of changes from London to East Anglia in England, suggests that the slow, unsalient, and phonetically gradual diffusion of fronter realizations of /ʌ/ ('cup' [kɐp – ka̠p]) is spreading in a wave-like way (1986: 51–53).

Urban Hierarchy Diffusion A much more common finding is an urban hierarchical effect, with innovations descending down a hierarchy of metropolis to city to town to village. One such example is highlighted by Kerswill's (2003: 235–236) investigations of the diffusion of labiodental variants of /θ ð/ in Britain. He showed that [f v], having diffused, he claims, from London, have reached the cities of Bristol, Derby, and Norwich *before* the nearer but smaller towns of Romsey and Wisbech (see also Chambers and Trudgill, 1998; Hernández Campoy, 2003). The usual explanation for the hierarchical model is that, while the distance between the diffusing center and the receiving location plays some role, interaction between urban centers in modern societies is likely to be greater and, therefore, a more frequent and effective conduit for the transmission of new forms than between urban and rural. Transportation networks tend to link urban with urban, and economic and consumption infrastructures tend to be based in and oriented toward urban centers, with the ensuing consequences for employment, commuting, and leisure patterns, and these feed the hierarchical nature of diffusion.

In research on the vocalization of /l/ in Australian English, Horvath and Horvath (1997) showed that a combination of the contagion and urban hierarchy models can help explain the diffusion of some changes. They found, quite unexpectedly, that rather

than being most prevalent in the larger cities of Sydney and Melbourne, /l/ vocalization was most vigorous in "the most slowly growing parts of the older core" (1997: 120) of the country, the South Australian centers of Adelaide and Mount Gambier. They propose a 'cultural hearth model' whereby an innovation gains a foothold in both town and country in one particular region before diffusing to other regions.

Contrahierarchical Diffusion Only very occasionally, innovations diffuse against the urban hierarchy. Trudgill (1986) described how a number of smoothing processes found in rural north Norfolk in eastern England were diffusing southward to urban centers in the county of Suffolk. Hence, 'power' was undergoing change from /pɑuə/ to /pɑː/, 'do it' from /dʉːəʔ/ to /dɜːʔ/ and 'sure' from /ʃʉːə/ to /ʃɜː/, etc.

The literature to date has not, broadly speaking, fully taken on board the consequences of seeing innovation diffusion as dialect contact – the contest between an incoming form and a local form – nor fully appreciated the problems with the approaches sociolinguists currently adopt with respect to diffusion, approaches which were abandoned in the human geographic literature over a quarter of a century ago (see Gregory, 1985, and for a sociolinguistic consideration of his arguments, Britain, 2002, 2005). Much of the literature, for example, assumes that, in diffusing spatially, the innovative form totally levels away the formerly used variants in its path, and that nonadoption of the innovation is a "passive state where the friction of distance applies a brake to innovation" (Gregory, 1985: 319). Furthermore, sociolinguistic models have concentrated on how the innovation spatially spreads almost to the exclusion of examining:

1. what happens to the innovation en route (Are the linguistic constraints on the use of the variant the same? Is the innovation evaluated in the same way in its heartland as it is in its new destination?); or
2. the social embedding of its promoters and receivers; or
3. the structural incongruities of the innovation and the traditional form it is attempting to eradicate.

In an investigation of the success (or otherwise) of a number of innovations spreading from London and the South-East of England to the Fens in the northwest of East Anglia, I showed (Britain, 2005) that, while some new forms were spreading successfully (all of which are notably less 'marked' than the forms they were replacing – e.g., the vocalization of /l/, the labiodentalization of prevocalic /r/, the fronting of /θ ð/ to [f v]), other 'innovations' were resisted (e.g., the fronting of /ʌ/). Furthermore, yet other innovations were only partly adopted, with the outcomes of the contact between old and new demonstrating reallocation or interdialect formation. In the East of the Fens, for example, the fronting of /ou/ (common in many dialects of English around the world) was only found in a historically defined subset of relevant lexical items – those deriving from Middle English **ou** (e.g., 'rows,' 'grown,' 'mown') – and was much less advanced in those words deriving from ME **oː** (e.g., 'rose,' 'groan,' 'moan'). How long this resistance to the innovation will last is questionable, but nevertheless the study shows that diffusion can lead to outcomes other than surrender by the local form (see also Trudgill, 1986; Britain, 2002). Hence we can see that innovation diffusion can be conceptualized in terms of the processes typical of contact – when an innovation meets a local form, koineization can occur (the mixture of variants tends to be 'resolved' in some way), and the outcomes can be leveling (one form wins over another) or reallocation/interdialect formation.

See also: Chinese; Dialect Chains; Falkland Islands: Language Situation; Language and Dialect: Linguistic Varieties; Pidgins and Creoles: Overview; Sociolect/Social Class; Sociophonetics; Spatial Variation (Geolinguistics); Spoken Discourse: Diachronic Aspects; Variation and Language: Overview.

Bibliography

Al-Dashti A (1998). 'Language choice in the state of Kuwait: a sociolinguistic investigation.' Unpublished Ph.D. dissertation. Colchester: Essex University.

Barz R & Siegel J (eds.) (1988). *Language transplanted: the development of overseas Hindi.* Wiesbaden: Otto Harrassowitz.

Britain D (1997). 'Dialect contact and phonological reallocation: 'Canadian Raising' in the English Fens.' *Language in Society 26,* 15–46.

Britain D (2001). 'If A changes to B, make sure A exists: a case study on the dialect origins of New Zealand English.' *Essex Research Reports in Linguistics 38,* 39–79.

Britain D (2002). 'Space and spatial diffusion.' In Chambers J K, Trudgill P & Schilling-Estes N (eds.) *Handbook of language variation and change.* Oxford: Blackwell. 603–637.

Britain D (2005). 'Innovation diffusion, 'Estuary English' and local dialect differentiation: the survival of Fenland Englishes.' *Linguistics 43.*

Chambers J (1992). 'Dialect acquisition.' *Language 68,* 673–705.

Chambers J & Trudgill P (1998). *Dialectology* (2nd edn.). Cambridge: Cambridge University Press.

Chou L (1997). 'Acquisition of lexical and phonological variants by Taiwanese students in England.' Unpublished M.A. dissertation. Colchester: Essex University.

Foreman A (2003). 'Pretending to be someone you're not: a study of second dialect acquisition in Australia.' Unpublished Ph.D. dissertation. Melbourne: Monash University.

Fox S (2005). 'The demise of 'Cockneys'? Language change in London's 'traditional' East End.' Unpublished Ph.D. dissertation. Colchester: University of Essex.

Gambhir S (1981). 'The East Indian speech community in Guyana: a sociolinguistic study with special reference to koine formation.' Unpublished Ph.D. dissertation. Philadephia: University of Pennsylvania.

Gregory D (1985). 'Suspended animation: the stasis of diffusion theory.' In Gregory D & Urry J (eds.) *Social relations and spatial structures*. London: Macmillan. 296–336.

Hernández Campoy J M (2003). 'Exposure to contact and the geographical adoption of standard features: two complementary approaches.' *Language in Society 32,* 227–255.

Horvath B (1985). *Variation in Australian English: the sociolects of Sydney.* Cambridge: Cambridge University Press.

Horvath B & Horvath R (1997). 'The geolinguistics of a sound change in progress: /l/ vocalisation in Australia.' In Boberg C, Meyerhoff M & Strassel S (eds.) *A Selection of Papers from NWAVE 25. Special issue of University of Pennsylvania Working Papers in Linguistics 4,* 109–124.

Kerswill P (1993). 'Rural dialect speakers in an urban speech community: the role of dialect contact in defining a sociolinguistic concept.' *International Journal of Applied Linguistics 3,* 33–56.

Kerswill P (1996). 'Children, adolescents and language change.' *Language Variation and Change 8,* 177–202.

Kerswill P (2002). 'Koineisation and accommodation.' In Chambers J K, Trudgill P & Schilling-Estes N (eds.) *Handbook of language variation and change*. Oxford: Blackwell. 669–702.

Kerswill P (2003). 'Dialect levelling and geographical diffusion in British English.' In Britain D & Cheshire J (eds.) *Social dialectology: In honour of Peter Trudgill.* Amsterdam: Benjamins. 223–244.

Khamkhong P (2004). 'Dialect contact and new dialect formation in a Thai New Town.' Unpublished Ph.D. dissertation. Colchester: University of Essex.

Kuo Y-H (2005). 'New dialect formation: the case of Taiwanese Mandarin.' Unpublished Ph.D. dissertation. Colchester: University of Essex.

Labov W (1966). *The social stratification of English in New York City.* Washington, DC: Center for Applied Linguistics.

Phili A (2003). 'How Famagustian do Famagustians sound? The presence of local variants in their speech, 30 years away from the town of Famagusta.' Unpublished M.A. dissertation. Colchester: Essex University.

Siegel J (1985). 'Koines and koineisation.' *Language in Society 14,* 357–378.

Sudbury A (2000). 'Dialect contact and koineisation in the Falkland Islands: development of a Southern Hemisphere English?' Unpublished Ph.D. dissertation. Colchester: Essex University.

Trudgill P (1986). *Dialects in contact.* Oxford: Blackwell.

Trudgill P (2004). *New dialect formation: the inevitability of colonial Englishes.* Edinburgh: Edinburgh University Press.

Watts E (2000). 'Acquisition of the Cheshire dialect by American expatriate children.' Unpublished M.A. dissertation. Colchester: Essex University.

Language: Mathematical Complexity

I Pratt-Hartmann, Manchester University, Manchester, UK

Basic Concepts

Let $f\colon \mathbb{N} \rightarrow \mathbb{N}$ be a function from natural numbers to natural numbers. Informally, an algorithm $\mathcal{A}$ has (time) complexity f if, given any input of size n, $\mathcal{A}$ will terminate after at most $f(n)$ steps. The subject of Complexity Theory is the mathematical study of the complexity of algorithms. This article explains the rudiments of Complexity Theory and outlines its application to the analysis of natural language.

Our first task is to make precise the concepts of 'algorithm,' 'size,' and 'step' mentioned in the previous paragraph. In the strictest sense, an *algorithm* is a deterministic Turing machine. Recall in this context that a *Turing machine* over a (finite) alphabet Σ comprises an infinite *tape* divided into discrete cells (or *squares*), a *head* that scans this tape one square at a time, and a finite table of *instructions*. Each instruction directs the head to write a symbol of Σ to the scanned square and then to move left or right, conditional on the current contents of the square and the current *machine state*. The number of possible machine states is finite, and the current machine state is updated, as each instruction is processed, according

Chambers J & Trudgill P (1998). *Dialectology* (2nd edn.). Cambridge: Cambridge University Press.
Chou L (1997). 'Acquisition of lexical and phonological variants by Taiwanese students in England.' Unpublished M.A. dissertation. Colchester: Essex University.
Foreman A (2003). 'Pretending to be someone you're not: a study of second dialect acquisition in Australia.' Unpublished Ph.D. dissertation. Melbourne: Monash University.
Fox S (2005). 'The demise of 'Cockneys'? Language change in London's 'traditional' East End.' Unpublished Ph.D. dissertation. Colchester: University of Essex.
Gambhir S (1981). 'The East Indian speech community in Guyana: a sociolinguistic study with special reference to koine formation.' Unpublished Ph.D. dissertation. Philadephia: University of Pennsylvania.
Gregory D (1985). 'Suspended animation: the stasis of diffusion theory.' In Gregory D & Urry J (eds.) *Social relations and spatial structures*. London: Macmillan. 296–336.
Hernández Campoy J M (2003). 'Exposure to contact and the geographical adoption of standard features: two complementary approaches.' *Language in Society 32,* 227–255.
Horvath B (1985). *Variation in Australian English: the sociolects of Sydney.* Cambridge: Cambridge University Press.
Horvath B & Horvath R (1997). 'The geolinguistics of a sound change in progress: /l/ vocalisation in Australia.' In Boberg C, Meyerhoff M & Strassel S (eds.) *A Selection of Papers from NWAVE 25. Special issue of University of Pennsylvania Working Papers in Linguistics 4,* 109–124.
Kerswill P (1993). 'Rural dialect speakers in an urban speech community: the role of dialect contact in defining a sociolinguistic concept.' *International Journal of Applied Linguistics 3,* 33–56.
Kerswill P (1996). 'Children, adolescents and language change.' *Language Variation and Change 8,* 177–202.
Kerswill P (2002). 'Koineisation and accommodation.' In Chambers J K, Trudgill P & Schilling-Estes N (eds.) *Handbook of language variation and change.* Oxford: Blackwell. 669–702.
Kerswill P (2003). 'Dialect levelling and geographical diffusion in British English.' In Britain D & Cheshire J (eds.) *Social dialectology: In honour of Peter Trudgill.* Amsterdam: Benjamins. 223–244.
Khamkhong P (2004). 'Dialect contact and new dialect formation in a Thai New Town.' Unpublished Ph.D. dissertation. Colchester: University of Essex.
Kuo Y-H (2005). 'New dialect formation: the case of Taiwanese Mandarin.' Unpublished Ph.D. dissertation. Colchester: University of Essex.
Labov W (1966). *The social stratification of English in New York City.* Washington, DC: Center for Applied Linguistics.
Phili A (2003). 'How Famagustian do Famagustians sound? The presence of local variants in their speech, 30 years away from the town of Famagusta.' Unpublished M.A. dissertation. Colchester: Essex University.
Siegel J (1985). 'Koines and koineisation.' *Language in Society 14,* 357–378.
Sudbury A (2000). 'Dialect contact and koineisation in the Falkland Islands: development of a Southern Hemisphere English?' Unpublished Ph.D. dissertation. Colchester: Essex University.
Trudgill P (1986). *Dialects in contact.* Oxford: Blackwell.
Trudgill P (2004). *New dialect formation: the inevitability of colonial Englishes.* Edinburgh: Edinburgh University Press.
Watts E (2000). 'Acquisition of the Cheshire dialect by American expatriate children.' Unpublished M.A. dissertation. Colchester: Essex University.

Language: Mathematical Complexity

I Pratt-Hartmann, Manchester University, Manchester, UK

Basic Concepts

Let f: $\mathbb{N} \rightarrow \mathbb{N}$ be a function from natural numbers to natural numbers. Informally, an algorithm $\mathcal{A}$ has (time) complexity f if, given any input of size n, $\mathcal{A}$ will terminate after at most $f(n)$ steps. The subject of Complexity Theory is the mathematical study of the complexity of algorithms. This article explains the rudiments of Complexity Theory and outlines its application to the analysis of natural language.

Our first task is to make precise the concepts of 'algorithm,' 'size,' and 'step' mentioned in the previous paragraph. In the strictest sense, an *algorithm* is a deterministic Turing machine. Recall in this context that a *Turing machine* over a (finite) alphabet Σ comprises an infinite *tape* divided into discrete cells (or *squares*), a *head* that scans this tape one square at a time, and a finite table of *instructions*. Each instruction directs the head to write a symbol of Σ to the scanned square and then to move left or right, conditional on the current contents of the square and the current *machine state*. The number of possible machine states is finite, and the current machine state is updated, as each instruction is processed, according

to that instruction. Computation proceeds as long as some instruction can be executed; if a point is reached where no instruction can be executed, computation terminates. A Turing machine is *deterministic* if, at any point in its computation, exactly one instruction may be executed. Thus, for a deterministic Turing machine $\mathcal{A}$, the entire course of the computation is determined by the string of symbols initially written on the tape.

If Σ is a finite alphabet of symbols, it is customary to denote the set of finite strings over Σ by Σ^*. Let $\mathcal{A}$ be a deterministic Turing machine over some finite alphabet Σ; we say that $\mathcal{A}$ *computes* the partial function $F: \Sigma^* \rightarrow \Sigma^*$ defined as follows: if $w \in \Sigma^*$ is written on the tape and the Turing machine run, then $F(w)$ is the string left on the tape when the computation terminates. (We take $F(w)$ to be undefined if the computation does not terminate.) The *size* of an input string w is simply its length and the *steps* of computation are simply the basic instructions as specified in $\mathcal{A}$'s instruction table. Thus, we can make precise the informal definition of algorithmic complexity with which we started: a deterministic Turing Machine $\mathcal{A}$ has time complexity f if, given any input string of size n, $\mathcal{A}$ will terminate after at most $f(n)$ steps. Of course, $\mathcal{A}$ will not in general take the same number of steps for all inputs of size n; the function $f(n)$ is thus an *upper bound* on running time. This kind of complexity measure – known as *worst-case* complexity – is the most common measure employed in Complexity Theory.

In measuring the complexity of an algorithm, our primary concern is with *asymptotic* behavior. Thus, it is considered significant whether an algorithm has complexity n^2 or 2^n, or even whether it is has complexity n^2 or n^3, but not whether it has complexity n^2 or $300\,n^2$. The motivation here is that linear factors can always be mitigated by running the algorithm on faster hardware. For this reason, the following terminology has become current. For any function, $f: \mathbb{N} \rightarrow \mathbb{N}$, we write $O(f(n))$ for the set of functions

$$\{t : \mathbb{N} \rightarrow \mathbb{N} \mid \exists\, c, m \in \mathbb{N} \textit{ such that } \forall n > m, t(n) \leq cf(n)\}.$$

That is, $O(f(n))$ is the set of functions from natural numbers to natural numbers that are eventually dominated by some fixed multiple of f. An algorithm $\mathcal{A}$ is then said to have (time) complexity $O(f(n))$ if it has (time) complexity t for some $t \in O(f(n))$. Beware: this an *asymptotic* measure. If one algorithm $\mathcal{A}$ has complexity $O(n^2)$, and another algorithm $\mathcal{A}'$ complexity $O(2^n)$, that does ***not*** mean that $\mathcal{A}$ need run faster than $\mathcal{A}'$ in practice; it merely means that, if the size of the inputs can increase without limit, a point will be reached at which $\mathcal{A}'$ takes longer than $\mathcal{A}$.

Thus far, we have discussed the complexity of *algorithms*, but often what is of interest is the complexity of the *functions* that those algorithms compute. And this notion too can be reconstructed using the apparatus introduced above. Let Σ be a finite alphabet, let $F: \Sigma^* \rightarrow \Sigma^*$ be a function from strings over Σ to strings over Σ, and let $f: \mathbb{N} \rightarrow \mathbb{N}$ be a function from natural numbers to natural numbers. Then we say that F has (time) complexity $O(f)$ if there exists a deterministic Turing machine $\mathcal{A}$ computing F such that $\mathcal{A}$ has time complexity $O(f)$. More generally, if $\boldsymbol{f}$ is a set of functions from natural numbers to natural numbers, we say that F is in $(\boldsymbol{f})$-**TIME** if there exists a Turing machine $\mathcal{A}$ computing F that terminates in a number of steps bounded by some function in $\boldsymbol{f}$. Consider, for instance, the set

$$\mathbf{P} = \{p(n) \mid p(n) \text{ a polynomial}\},$$

giving us the complexity class **PTIME**. To say that a function F is in **PTIME** – or simply, 'is polynomial' – is to say that it can be computed by a deterministic Turing machine guaranteed to halt in a number of steps bounded by some polynomial function of the size of the input. Other salient sets of functions are

$$\mathbf{EXP} = \left\{2^{p(n)} | p(n) \text{ a polynomial}\right\}$$
$$\mathbf{2\text{-}EXP} = \left\{2^{2^{p(n)}} | p(n) \text{ a polynomial}\right\},$$

and so on. These give us, respectively, the computational complexity classes **EXPTIME**, **2-EXPTIME**, and so on. The union of the complexity classes ***K*-EXPTIME** for all $K \in \mathbb{N}$ is called the class of *elementary* functions. Functions F lying outside this class are called *nonelementary.* Finally, we note that, for a given function $F: \Sigma^* \rightarrow \Sigma^*$, it is not necessarily the case that there ***is*** a Turing machine that computes F at all. Such functions are said to be *noncomputable*.

Of particular interest in Complexity Theory are functions $F: \Sigma^* \rightarrow \{0,1\}$. Such functions are known as *decision problems*. (The intuition is that $F(w) = 1$ represents the answer 'yes' for a given input w, and $F(w) = 0$ represents the answer 'no'.) The following important notions relate specifically to decision problems. Let $\mathcal{A}$ be a Turing machine (not necessarily deterministic) and $F: \Sigma^* \rightarrow \{0,1\}$ a decision problem. We say that $\mathcal{A}$ *nondeterministically computes* F if, for all input strings $w \in \Sigma^*$, we have $F(w) = 1$ if and only if there exists a run of $\mathcal{A}$ that terminates with 1 written on the tape. Finally, let $\boldsymbol{f}$ be a set of functions from natural numbers to natural numbers. We say that the decision problem F is in the class $\mathbf{N}(\boldsymbol{f})$-**TIME** if F is nondeterministically computed by a Turing machine that terminates in time bounded by a function in $\boldsymbol{f}$.

As with the deterministic case, an important special case is the set **P**, giving us the complexity class **NPTIME** (or simply **NP**, as it is traditionally called). Similarly, we have the complexity classes **NEXPTIME**, **2-NEXPTIME**, and so on. Note the asymmetry in the definition of nondeterministic computation: if $F(w) = 1$, then a Turing machine nondeterministically computing F must have at least one run on input w terminating with output 1; if $F(w) = 0$, by contrast, no such run exists. For this reason, we sometimes need to speak about the complementary complexity classes **Co-NP**, **Co-NEXPTIME** etc.; a decision problem F is said to be in the class **Co-NP** if the decision problem F defined by $\overline{F}(n) = 1 - F(n)$ is in **NP**, and so on.

Just as we can count the *steps* involved in a run of some Turing machine, so too can we count the number of *tape squares* that it accesses. Again, we typically specify an upper bound on this number as a function of the size of the input. These complexity measures lead to classes **PSPACE**, **EXPSPACE**, etc., defined analogously as for **PTIME**, **EXPTIME**, etc. It is not necessary to talk about nondeterministic space-complexity classes, since they can be shown to coincide with their deterministic counterparts. In fact, the complexity classes mentioned above satisfy the inclusions:

$$\mathbf{PTIME} \subseteq \mathbf{NP} \subseteq \mathbf{PSPACE} \subseteq \mathbf{EXPTIME} \subseteq \mathbf{NEXPTIME} \subseteq \mathbf{EXPSPACE} \subseteq \mathbf{2\text{-}EXPTIME} \subseteq \ldots \quad (1)$$

The inclusions **PTIME** $\subset$ **EXPTIME** $\subset$ **2-EXPTIME** ... are strict: there are problems in **EXPTIME** that are not in **PTIME**, problems in **2-EXPTIME** that are not in **EXPTIME**, and so on. Whether the other inclusions in (1) are strict is a fundamental unsolved problem.

There is one final topic to cover before we apply these ideas to natural language. Suppose that F and F' are decision problems, with the following property: there is a function $r: \Sigma^* \to \Sigma^*$, which can be computed in polynomial time, such that, for all strings w, $F(r(w)) = F'(w)$. In such a case, we say that F' is (polynomially) *reducible to F*. If F' is (polynomially) reducible to F, any method for solving the problem F automatically yields a method for solving the problem F': first apply the function r to translate any given instance of F' to a corresponding instance of F and then apply the method for solving F. The time taken to solve F' in this way is no greater than the time taken to solve F' plus some polynomially bounded number of steps for the translation. Suppose now that it is possible to establish a result of the following form concerning a decision problem F: for all problems F' in some complexity class, F' is polynomially reducible to F. We then say that F is **hard** for the complexity class in question. Thus, for example, a problem to which any problem in the class **NP** can be polynomially reduced is said to be **NP**-hard, and so on. Intuitively, a decision problem is hard for a complexity class if it is at least as difficult to solve as any problem in that complexity class. Of particular interest are those problems that are simultaneously in some complexity class and hard for that class. Such problems are said to be *complete* for the class in question. Thus, problems can be **NP**-complete, **PSPACE**-complete, **EXPTIME**-complete, etc.

The theoretical framework presented thus far has employed the standard – though somewhat austere – view of an algorithm as a Turing machine operating on strings of symbols taken from a finite alphabet. In practice, however, when analyzing the complexity of a specific algorithm or function, a more relaxed concept of algorithm and a richer collection of data types are employed. There is no one 'correct' notion of basic operation here, but the fundamental principle is that basic operations should be executable in an amount of time that is, for the purposes at hand, bounded by a constant. Thus, for example, in problems where all the integers encountered are likely to remain small, it may be reasonable to regard the addition of two integers as a basic operation, since doing so gives a good estimate of actual running time. Similar remarks apply to measuring the size of the various data structures constituting the inputs of an algorithm.

Sometimes, the size of an input is best represented by more than one parameter, especially if that input is a tuple of data structures. The O-notation introduced above is extended to cover this case in the obvious way. For example, let $f: \mathbb{N}^2 \to \mathbb{N}$ be a function from pairs of natural numbers to natural numbers. We write $O(f(m,n))$ for the set of functions

$$\{t : \mathbb{N}^2 \to |\exists c, l \in \mathbb{N} \text{ such that } \forall n, m > l, t(m,n) \leq cf(m,n)\}.$$

An algorithm is then said to have (time) complexity $O(f(m,n))$ if its running time for any input of size m, n is bounded by some function $t(m,n) \in O(f(m,n))$.

For a comprehensive introduction to Turing machines and Complexity Theory, we recommend Papadimitriou (1994).

Recognizing Sentences

Let us see how the ideas sketched in the previous section can be applied to problems in linguistics. We begin with the recognition problem for context-free

grammars. Formally, a *context-free grammar* (CFG) is a quadruple $G = \langle N, V, S, P\rangle$, where N is a set of *nonterminals* (typically, category labels such as S, NP, VP, etc.), V is a set of terminals (typically, words such as *every*, *boy*, *loves*, etc.), S is a distinguished *start symbol* in N (for example, the category S) and P is a list of *productions* for rewriting nonterminals (such as, S → NP VP, NP → Det N, etc.). A CFG *accepts* the string of terminals w if some sequence of productions can be found that rewrites the start symbol S to w. (For a detailed discussion, see, e.g., Hopcroft and Ullman, 1979: 77 ff.) The *language of* G, denoted $L(G)$, is the set of strings accepted by G. The recognition problem for CFGs is: given a CFG G and a string w, is w a member of $L(G)$?

For simplicity, we assume that G is in Chomsky normal form; that is, all productions in G are of the forms $X \to YZ$ or $X \to a$, where X, Y, and Z are nonterminals and a is a terminal. (This restriction does not reduce the expressive power of the formalism.) Under this assumption, the algorithm of Figure 1 solves the CFG recognition problem for the grammar $G = \langle N, V, S, P\rangle$, and string $w = w[0]\ldots w[n-1]$. This algorithm is generally known as the CYK algorithm after its discoverers, Cocke, Younger, and Kasami (see, e.g., Younger, 1967). The basic idea is to maintain an array M (usually called a *chart*) such that $M[i][j]$ stores the grammatical categories that G can assign to the substring $w[i]\ldots w[j-1]$. The loops in lines 2–6 record the possible categories of each word (i.e., each 1-element substring) of w, and the loops in lines 7–16 record the possible categories of successively longer substrings of w.

Let us suppose that the individual instructions in Figure 1 denote basic operations. The complexity of this algorithm is then easily seen to be $O(mn^3)$, where m is the number of productions in G, and n is the length of w. Hence, the general recognition problem for CFGs has complexity $O(mn^3)$; moreover, for every fixed CFG G, the problem of determining whether a string w is in $L(G)$ has complexity $O(n^3)$. We note in passing that our choice of basic operations in this analysis is not entirely innocent. For example, the tests in line 11 read from a list of up to m elements, and the instruction in line 12 writes to such a list. In practice, however, these operations are so fast that it is reasonable to pretend that they are performed in constant time.

```
1 begin cyk(G, w)
2   for i = 0 to n–1
3     for each A → a ∈ P
4       if w[i] = a then add A to M[i][i+1]
5       endfor
6   endfor
7   for i = 0 to n–1
8     for j = i + 2 to n
9       for each A → BC ∈ P
10        for k = i + 1 to j–1
11          if B ∈ M[i][k] and C ∈ M[k][j] then
12          add A to M[i][j]
13        endfor
14      endfor
15    endfor
16 endfor
17 if S ∈ M[0][n] then return 1
18 else return 0
19 end cyk
```

Figure 1 The CYK algorithm.

Of course, CFGs are not the only way of describing languages. One illustrative alternative is provided by the linguistic formalism known as *Lambek grammar*. The reader is referred to Carpenter (1998) for details; we content ourselves with an informal explanation here. The *Lambek calculus* is a logical system allowing the derivation of various *sequents* involving *category expressions*. For example, Figure 2 shows a derivation of the sequent NP (NP\S)/NP NP → S. Each step in this derivation must be licensed by one of a small collection of rules of inference. Intuitively, a category expression X/Y describes a string that, when a string of category Y is placed to its *right*, will result in a string of category X; similarly, $Y\backslash X$ describes a string that, when a string of category Y is placed to its *left*, will result in a string of category X. Under this interpretation, a sequent $X_0 \ldots, X_{n-1} \to X$ has the meaning: 'Any string of items of categories $X_0 \ldots X_{n-1}$ is of category X.' Now, a *Lambek grammar* G is a finite list of pairs of the form (v, C), where v is a vocabulary item and C is a category expression. We say that the grammar G *accepts* the string $w = w[0] \ldots w[n-1]$ just in case there exist $C_0 \ldots, C_{n-1}$ such that: (i) $(w[i], C_i) \in G$ for each i $(0 \le i < n)$ and (ii) the sequent $C_0 \ldots C_{n-1} \to S$ can be derived in the Lambek calculus. (Again, S is a distinguished start symbol.) Suppose, for example, G contains the pairs (*John*, NP), (*Mary*, NP) and (*loves* (NP\S)/NP). Since NP (NP\S)/NP NP → S is a derivable sequent, it follows that G accepts the sentence *John loves Mary*. The *language of* G, denoted $L(G)$, is the set of strings accepted by G. The recognition problem for Lambek grammars is as follows: given a Lambek grammar G and a string w, is w a member of $L(G)$?

Since the category expressions in a Lambek grammar may be complex, a reasonable measure of the size

$$\dfrac{\dfrac{\mathrm{S} \to \mathrm{S} \qquad \mathrm{NP} \to \mathrm{NP}}{\mathrm{NP}\;\;(\mathrm{NP}\backslash\mathrm{S}) \to \mathrm{S}} \qquad \mathrm{NP} \to \mathrm{NP}}{\mathrm{NP}\;\;(\mathrm{NP/S})\,/\,\mathrm{NP}\;\;\;\mathrm{NP} \to \mathrm{S}}$$

Figure 2 A derivation in the Lambek calculus.

m of G is in this case the total number of atomic symbols involved in a complete listing of G. Thus, we may reasonably take the total size of the input to the recognition problem to be $m+n$, where n is the length of w. On this measure of input size, it follows from the cut-elimination theorem for the Lambek calculus (Lambek, (1958)) that the recognition problem for Lambek grammars is in the complexity class **NP**. Furthermore, it has been conjectured that this problem is **NP**-hard. If so, it may well be that, unlike the recognition problem for CFGs, no polynomial-time algorithm exists to solve it.

It is known (Pentus, 1993, 1997) that the class of languages recognized by Lambek grammars is identical to the class of languages recognized by context-free grammars. Yet, as we have seen, the recognition problem for CFGs is $O(mn^3)$. This illustrates that the complexity of the recognition problem for a particular grammar formalism cannot be read off in any simple way from its expressive power: CFGs and Lambek grammars have the same expressive power, but their recognition problems may have different complexity. Of course, the expressive equivalence of CFGs and Lambek grammars does tell us that, for a *fixed* Lambek grammar G, the complexity of determining whether a given string (of length n) is in $L(G)$ is $O(n^3)$. (Just get a CFG recognizing the same language as G and use the CYK algorithm!)

These cautionary remarks notwithstanding, commonly encountered grammar formalisms with higher expressive power do tend, unsurprisingly, to exhibit higher recognition complexity. A well-known example is provided by the class of *tree adjoining grammars* (TAGs). A comprehensive introduction to TAGs can be found in Abeillé and Rambow (2000: Chap. 1). But very roughly, a TAG is a finite set of 'local' trees that can be combined into larger trees to license sentences, much as CFGs combine productions (which can equally be thought of as local trees) into phrase structures. The essentially new element in TAGs is the operation of *adjunction*, in which a local tree may be 'spliced' into an existing tree. This leads to an increase in recognition capacity: there exist languages that are recognized by some TAG, but not by any CFG. It also apparently leads to an increase in recognition complexity: the best known parsing algorithms for TAGs have time complexity $O(n^6)$, where n is the length of the input string (Schabes, 1994).

Moving up the expressiveness hierarchy, a *context-sensitive grammar* (CSG) is a set of productions $\alpha \to \beta$, where α and β are strings of symbols such that $|\alpha| \leq |\beta|$. These productions are interpreted as rewrite rules, in much the same way as productions of a CFG. For comparison, note that, in a CFG, all productions have the form $A \to \beta$, where A is a nonterminal. Indeed, if we assume (which we may without essential loss of expressive power) that productions in CFGs have nonempty right-hand sides, the condition $|\alpha| \leq |\beta|$ is then trivially satisfied, whence CFGs are a special case of CSGs. The *language* $L(G)$ *of* a CSG G is defined analogously to the context-free case. It turns out that the problem of deciding whether, for a given context-sensitive grammar G and a given string w, w is in $L(G)$ is **PSPACE**-complete. (In fact, it is possible to find a *fixed* context-sensitive grammar G for which the problem of determining whether a given string w is in $L(G)$ is PSPACE-complete.) For a formal definition of context-sensitive grammar and a proof of this complexity result, see Hopcroft and Ullman (1979: 223, 347 ff.).

All the grammar formalisms examined thus far have precise formal definitions, which makes for a clear-cut (though often nontrivial) complexity analysis. However, many mainstream grammar formalisms that aspire to describe natural languages are much less rigidly defined (and indeed much more liable to periodic revision); consequently, it is harder to provide definitive results about computational complexity. Transformational grammar, as outlined in the *Aspects* theory of Chomsky (1965), provides a case in point. Let us take a transformational grammar to consist of two components: a context-free grammar generating a collection of phrase-structure trees – so-called *deep structures* – and a collection of *transformations* that map these deep structures to other phrase-structure trees – so-called *surface structures*. The movement principles in question are too complex for a useful discussion here. The *language of* a transformational grammar G, denoted $L(G)$, can then be defined as the set of strings w such that w is the string of leaves of some surface structure obtained in this way. Analyzing a version of Chomsky's *Aspects* theory, Peters and Ritchie (1973) demonstrated the existence of transformational grammars that can recognize nonrecursive sets. That is to say, the recognition problem for transformational grammars (thus understood) is not even computable. More recent analyses of grammar formalisms in the transformational tradition paint a picture of lower complexity, however. Thus, Berwick and Weinberg (1984: 125 ff.) analyze the complexity of *Government-Binding Grammars*, a formalization of the approach taken in Chomsky (1981), and showed that the recognition problem for such grammars is in PSPACE.

Recent trends in linguistics – particularly within the transformational tradition – have shown a preference for specifying grammars not in terms of generative mechanisms, but rather, in terms of constraints to which sentence structures are required to conform. On this view of grammar, a string w is grammatical

just in case it has a structure that satisfies those constraints. How can we determine the complexity of the recognition problem when grammars are presented in this way? The answer is to employ a formal language: this formal language must be powerful enough to express the constraints constituting the grammar in question, and yet not so powerful that working with it leads to undecidable problems. *Model theoretic syntax* (Rogers, 2003) is a grammar formalism employing this strategy. To understand how it works, we must first rehearse some ideas from formal logic.

Weak monadic second-order logic (wMSO) employs a formal language containing two sorts of variables: those ranging over objects (as in ordinary first-order logic) and those ranging over *finite sets* of objects. For present purposes, we may assume that the 'objects' in question are nodes in a tree that describes the phrase structure of some sentence. We confine ourselves to a language containing binary atomic formulas of the forms $x \in X$, $x \lhd_1 y$ and $x \lhd_2 y$, and unary atomic formulas of the forms $np(x)$, $vp(x)$, etc., where *np*, *vp*, etc., form a finite collection of labels. Consider a tree in which each node is tagged with one of these labels. Then the above atomic formulas may be interpreted as follows: $x \in X$ means that node x is a member of the set of nodes X; $x \lhd_1 y$ means that node x is the left sister of node y; $x \lhd_2 y$ means that node x is the mother of node y; $np(x)$ means that x is tagged with the label *np*; and so on. Formulas are built up from atomic formulas using Boolean connectives and quantifiers (over both sorts of variables) in the normal way. The semantics given above then determine, for a given closed formula and a given labeled tree, whether that formula is true in that labeled tree.

On this view, we can think of a grammar G as a closed formula of weak monadic second-order logic. The set of trees recognized by G is simply the set of trees in which G is true; the language $L(G)$ recognized by G can then be defined as the set of strings w such that w is the string of leaves of some tree recognized by G. The appeal of this approach is twofold. First, wMSO can express many relationships dear to linguists' hearts. For example, it is straightforward to write down a formula $\phi_c(x, y)$, which is satisfied by nodes x and y in a tree just in case x C-commands y. Second, despite its high expressive power, this language is decidable. That is, given any closed formula ϕ, we can effectively determine whether there is a labeled tree in which it is true. (This follows from Rabin's theorem on the decidability of the monadic second-order theory of two successor functions.) It follows that, given any grammar G and any string w, we can effectively determine whether $w \in L(G)$. As might be expected for such a general approach, the complexity of such model theoretic grammars is relatively high. In fact, the problem of deciding satisfiability in this logic is nonelementary. (See Börger *et al.*, 1997: 315 ff., for an accessible treatment.) This complexity result implies that this way of defining grammars is not expressive enough to capture the transformational grammars analyzed by Peters and Ritchie. However, more recent theories in the same tradition may be amenable to formalization within this framework. For example, Rogers (2003) noted that some principles of Rizzi's theory of *Relativized Minimality* (Rizzi, 1990) can be expressed using formulas of the language sketched above.

Recovering Meaning

Most grammars are more than a means of defining a set of acceptable sentences: they also assign one or more levels of structure to those sentences that they do accept. The question then arises as to the computational complexity of recovering that structure.

Consider, for example, context-free grammars. Let G be a CFG. If $w \in L(G)$, then G assigns to w one or more phrase structures representing the derivation of w by the productions of G. It is easy to construct a CFG G for which there exists a sequence $\{w_n\}$ of strings accepted by G, such that the length of w_n is bounded above by some polynomial function of n while the number of phrase structures that G assigns to w_n is bounded below by an exponential function of n. That is, the number of parses produced by a CFG G can grow exponentially. Nevertheless, the CYK algorithm of Figure 1 shows that the set of phrase structures assigned to any sentence by G may always be compactly represented; moreover, that compact representation may be constructed in time $O(mn^3)$. To see this, note that the chart M in this algorithm specifies every category that G assigns to every substring of w. It is easy to modify the data structures stored in the entries of M so that the elements of $M[i, j]$ are not simply category labels, but rather category labels with pointers to entries in $M[i, k]$ and $M[k, j]$ indicating their daughters. The result will be an acyclic directed graph that can be expanded into a complete list of all the parse trees assigned to w by G in the obvious way. (Trivially, listing all these trees will in general take exponential time.)

Arguably, determining the syntactic structure of a sentence is of little value unless we can use that structure to recover the sentence's meaning. The notion of meaning in general is too vague to admit of immediate formal analysis. However, we might sensibly

begin with the more specific problem of recovering, at least for certain fragments of natural languages, *logical form*, in the sense of producing translations such as the following:

Every boy loves some girl who admires him.

$$\forall x(\text{boy}(x) \rightarrow \exists y(\text{girl}(y) \wedge \text{admire}(y, x) \wedge \text{love}(x, y))) \qquad (2)$$

Context-free grammars (and indeed the other grammar formalisms mentioned above) can be modified to yield such logical forms. Approaches vary, but one commonly employed technique is to associate with each vocabulary item an expression of higher-order logic representing its meaning and to associate with each production a prescription for combining the meanings of its daughters. This approach is, more or less, that championed by Montague (1974); for an accessible modern treatment, including an explanation of the relevant principles of higher-order logic, see, for example, Heim and Kratzer (1998). The question now arises: what is the computational complexity of recovering logical forms in this way?

The answer depends on how exactly logical forms are allowed to be represented. If the underlying grammar G is a CFG, then the CYK algorithm can again be modified to produce, in polynomial time, a compact representation of all meanings that G assigns to a given string w, just as for parse trees. However, these representations will not, in the parlance of higher-order logic, be 'β-reduced.' That is, in order to produce ordinary logical translations such as (2), we need to compute the normal forms for the expressions of higher-order logic that our parser yields. And this problem is known to be nonelementary (Statman, 1979). It follows that semantically annotated CFGs can be written for which the problem of computing normalized translations is also nonelementary. In practice, however, the normalization of semantic representations produced by semantically annotated CFGs never constitutes a bottleneck; for most grammars in this form that one is likely to write, computing the semantics entails no essential change in computational complexity. At present, however, there is no systematic formal analysis that can be used to underpin this claim.

Working with Meanings

Finally, even when sentences have been parsed and meanings recovered, the question arises as to the complexity of *using* these sentences. The uses to which language is put are too varied for this problem to admit of any immediate formal analysis. However, we might sensibly begin with the following, more specific problem: given a grammar describing a fragment of natural language, what is the complexity of determining valid entailments involving sentences of that fragment?

A familiar example should help motivate this problem. Consider the set of all English sentences subsumed by the following six schemata, where A and B are any common (count) nouns and S is any proper noun:

Every A is a B.	No A is a B.
Some A is a B.	Some A is not a B.
S is a B.	S is not a B.

Since all sentences in this fragment use the copula to relate two terms, let us call the fragment *Cop*. Cop is, of course, just the language of the Syllogistic. Since the set of sentence types in Cop is finite, the translation from its sentences to first-order formulas representing their logical forms certainly takes only polynomial time. Let e be a sentence in Cop, and let E be a set of such sentences. Let us say that E *entails* e just in case the logical translations of E entail the logical translation of e (in the usual sense of first-order logic). Then we ask: what is the complexity of determining, given E and e, whether E entails e? It is easy to show using standard techniques in logic that this problem is in fact $O(n^2)$, where n is the total number of words in $E \cup \{e\}$.

Now suppose we add relative clauses to Cop, generating such sentences as, for example,

Every philosopher who is not a stoic is a cynic.

Let us call this larger fragment *Cop + Rel*. It is simple to specify Cop + Rel in terms of a grammar belonging to any of the grammar formalisms discussed above. Furthermore, by equipping the grammar in question with a semantics, as discussed in the previous section, we obtain a (polynomial-time) algorithm translating sentences of Cop + Rel into first-order logic, thus defining a notion of entailment exactly as for the fragment Cop. Again we ask: what is the complexity of the entailment problem for Cop + Rel? The answer this time is that it is **Co-NP**-complete. Thus, relative clauses may increase the complexity of determining entailments. Now suppose that we add constructions involving transitive verbs to Cop + Rel, generating sentences such as, for example:

No philosopher who despises some stoic admires every cynic.

Call the resulting fragment *Cop + Rel + TV*. What is the complexity of determining validity now? This

time, it has increased: the validity problem for Cop + Rel + TV is **EXPTIME**-complete.

Likewise, suppose that we add constructions involving ditransitive verbs to Cop + Rel + TV, generating sentences such as, for example,

> No philosopher who despises some stoic recommends every cynic to every skeptic.

Call the resulting fragment *Cop + Rel + TV + DTV.* The validity problem for Cop + Rel + TV + DTV, it transpires, is **Co-NEXPTIME**-complete. Alternatively, suppose we add anaphora to Cop + Rel + TV, thus giving us the ability to recover sentence (2) above. Here, the situation is more complex, because of the need to deal with anaphoric ambiguities. However, it can be shown that, if anaphors are required to take as antecedents the closest possible noun phrase (consistent with the normal rules of binding theory), the result is a fragment in which the validity problem is **Co-NEXPTIME**-complete. Relaxing this restriction results in a fragment for which the validity problem is undecidable.

To be sure, determining entailments between sentences is a rather specialized, not to say abstruse, activity, far removed from the cut-and-thrust of everyday language use. However, the analysis sketched above illustrates some of the broader issues that arise in connection with the computational complexity of natural language. For there is surely more to human beings' considerable linguistic prowess than their ability to recognize membership in sets of strings. Technical details can be found in Pratt-Hartmann (2004).

See also: Chomsky, Noam (b. 1928); Montague, Richard (1931–1971); Montague Semantics.

Bibliography

Abeillé A & Rambow O (2000). *Tree adjoining grammars: Formalisms, linguistic analysis and processing.* Stanford, CA: CSLI Publications.

Berwick R C & Weinberg A S (1984). *The grammatical basis of linguistic performance: Language use and acquisition.* Cambridge, MA: MIT Press.

Börger E, Grädel E & Gurevich Y (1997). *The classical decision problem.* Perspectives in Mathematical Logic. Berlin: Springer-Verlag.

Carpenter B (1998). *Type-logical semantics.* Cambridge, MA: MIT Press.

Chomsky N (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Chomsky N (1981). *Lectures on government and binding.* Dordrecht: Foris Publications.

Heim I & Kratzer A (1998). *Semantics in generative grammar.* Oxford: Blackwell.

Hopcroft J E & Ullman J D (eds.) (1979). *Introduction to automata theory, languages and computation.* Reading, MA: Addison Wesley.

Lambek J (1958). 'The mathematics of sentence structure.' *American Mathematical Monthly 65(3),* 154–170.

Montague R (1974). 'The proper treatment of quantification in ordinary English.' In Thomason R (ed.) *Formal philosophy.* New Haven, CT: Yale University Press. 247–270.

Papadimitriou H (1994). *Computational complexity.* Reading, MA: Addison-Wesley.

Pentus M (1993). 'Lambek grammars are context free.' In *Proceedings of the 8th Annual Symposium on Logic in Computer Science.* IEEE Press.

Pentus M (1997). 'Product-free Lambek calculus and context-free grammars.' *Journal of Symbolic Logic 62(2),* 648–660.

Peters P S Jr. & Ritchie R W (1973). 'On the generative power of transformational grammars.' *Information Sciences 6,* 49–83.

Pratt-Hartmann I (2004). 'Fragments of language.' *Journal of Logic, Language and Information 13,* 207–223.

Rizzi L (1990). *Relativized minimality.* Cambridge, MA: MIT Press.

Rogers J (2003). 'Syntactic structures as multi-dimensional trees.' *Research on Language and Computation 1(3/4),* 265–305.

Schabes Y (1994). 'Left-to-right parsing in lexicalized tree-adjoining grammars.' *Computational Intelligence 10(4),* 506–524.

Statman R (1979). 'The typed λ-calculus is not elementary recursive.' *Theoretical Computer Science 9,* 73–81.

Younger D H (1967). 'Recognition and parsing of context-free languages in time n^3.' *Information and Control 10,* 189–208.

Language: Semiotics

F Nuessel, University of Louisville, Louisville, KY, USA

Introduction

To discuss the semiotics of language entails a consideration of both of these terms. The person most consider to be one of the founders of semiotics, Charles Sanders Peirce (1839–1914), defined semiotics as 'the doctrine of signs.' Simply stated, a sign is something that stands for something else. The Greek philosopher, Aristotle (384–322 B.C.E.), provided a tripartite theory of signs. A sign thus has a physical dimension (a physical object), a referent (what the sign refers to), and an interpretant (what the sign means). Greek physicians, who were the first to discuss signs, used the notion of sign to refer to symptomatology. Certain physical signs stood for particular ailments (redness and swelling might signify a venomous animal bite).

For Sebeok (1994: 17–41), there are six types of signs: signals (the discharge of a pistol to start a race), symptoms (physical signs such as bruises), icons (images, diagrams, metaphors), indexes (weather cock, photograph), symbols (culturally determined meanings accorded to objects, e.g., a cross), and names (how we label ourselves and our physical world). In this article, we discuss several of these sign types as they relate to language.

Language refers to verbal communication. In this sense, it is distinct from nonverbal forms of communication such as kinesics (significant body movement), proxemics (personal space), paralanguage (vocal modification), and codes (Braille and the Morse code). Language thus enjoys a privileged place because it is species unique (*Homo sapiens*).

When we speak of the semiotics of language, we must consider its status with relation to semiotics. Language is a subdomain of semiotics because it is a form of communication. In this article, we look at language as a code that allows us to represent our physical world via linguistic signs (the sounds, words, and syntax of specific languages, and their corresponding written versions) to refer to external reality. What we encode in language reflects what is symbolically significant in a particular culture. Sometimes this symbolization is direct and obvious, i.e., we devise words to express these culturally significant concepts. In other cases, its more subtle meanings are encoded in other ways, e.g., words to indicate social relationships or word order to specify new versus presupposed information.

Language is ubiquitous. All of our daily activity involves language. We use language in all of our interpersonal relationships, the messages that we send to others electronically or in written form, the books that we read, the lyrics of the music we listen to, the advertising that bombards us continually, are but a few examples. Language is the mechanism that encodes culture, the sum total of a group's belief systems, values, and creative activity and it transmits this knowledge intergenerationally.

In this article, we discuss the following aspects of language:

1. A definition and discussion of the notion of language;
2. Linguistic theory;
3. The origination and loss of language(s);
4. Language change, or how language changes through time and space;
5. Pragmatics, or how language is used in different contexts;
6. Neurolinguistics, or the relation of the brain to language;
7. Language games, or the way we play with language;
8. The language of names;
9. The language of advertising.

What Is Language?

What is language? The answer to this fundamental question is not simple. A definition that professors of linguistics often use is the following one: language is an arbitrary system of vocal (or visual) symbols used by the members of a speech community for the purpose of communication. This definition consists of six elements that merit commentary here.

Language is arbitrary because the sounds and forms that different languages use to refer to objects, concepts, and activities are vastly different among the estimated 6000 languages of the world. The relationship of a word and the object or concept to which it refers is thus almost always arbitrary. Although arbitrary in its form–meaning relationship, language is subject to constraints. These controls include limits on the possible type and number of meaningful sounds permissible in a language (20–44 distinctive sounds appear to be the limit). Likewise, word order has limits, and there appear to be six possibilities. The formation of words is also subject to regulation in terms of permissible prefixes, suffixes, and infixes. Second, language is systematic because it employs a few basic rules to create an infinite number of possible combinations, e.g., the phonology (sounds), morphology (word formation rules), syntax (rules that determine the order and arrangement of words within

a sentence), lexicon (the vocabulary of a language), and semantics (the meanings of words in isolation and in context) to formulate meaningful interaction. Third, language is generally vocal, although certain languages used by the deaf (American Sign Language) employ visual cues. Fourth, language is symbolic because its sounds or visual emblems stand for ideas and concepts that reflect the organization of the external world. Fifth, the members of a speech community use a particular language to express their culture, i.e., its thoughts, ideas, and emotions. This community may be local, e.g., the immigrants from another country who retain their first language in order to maintain their identity. Alternatively, the community may be international, e.g., the worldwide community of speakers of Spanish, Italian, or English. Finally, the speech community uses a language to express its feelings, beliefs, and emotions to other speakers of the language.

It should be noted that there is a use of the word 'language' that is applied to certain creative activities. We often speak of the language of film, the language of art, the language of architecture, the language of music, and so forth. This usage is metaphoric because it means that artists employ certain conventions that may be equated with the grammatical rules of a language, i.e., rules for combining the elements of a particular artistic craft into meaningful compositions that may be interpreted by others.

Design Features of Language

Charles F. Hockett (1914–2000) set out to define and differentiate human language and animal communication. In his article on linguistic universals (Hockett, 1963: 8–13), Hockett specifies a set of 16 "design features" for comparing communication systems across species. The latter essay was an expansion of an earlier one in *Scientific American* (Hockett, 1960: 90) in which Hockett originally proposed thirteen design features. These design features provide an excellent overview of the basic elements of language. Table 1 illustrates this typology.

Table 1 Design features of language

Design feature	*Characteristics*
1. Auditory–vocal channel	This feature means that language is an oral–aural set of process rather than a visual, tactile, gustatory, or olfactory procedure.
2. Broadcast transmission and directional reception	This feature means that a verbal signal may be heard within the limits of the ear's ability to detect sound and determine its source.
3. Rapid fading (transitoriness)	This feature means that sounds are temporary and do not last.
4. Interchangeability	This feature means users of a language can duplicate any message they are capable of understanding.
5. Total feedback	This feature means that users of a language can reflect upon what they hear.
6. Specialization	This feature means that the sound waves of a language have the specific purpose of communicating meaning.
7. Semanticity	This feature means that linguistic components carry meaning because they have a systematic reference to objects in the world.
8. Arbitrariness	This feature means that the relationship of the referent and the sound employed to refer to it are independent.
9. Discreteness	This feature means that the sound system of a language consists of a small set of the possible sounds available to it and that these sounds have a pattern of meaningful oppositions, e.g., the consonants /s/, /z/ which are differentiated solely by voicing in the English language.
10. Displacement	This feature means that language may refer to circumstances that are temporally or geographically remote.
11. Openness/ Productivity	This feature means that languages use a finite set of elements (sounds, words, syntax) to create an infinite set of linguistic possibilities.
12. Traditional transmission	This feature means that language is transmitted in a given culture from one generation to another through accepted principles of teaching and learning.
13. Duality of patterning	This feature means that verbal sounds bear no meaning until they combine in a systematic way with other sounds to form meaningful linguistic components such as words and sentences that convey significance in a particular culture.
14. Prevarication	This features means that messages may be false or logically meaningless.
15. Reflexiveness	This feature refers to metalanguage, the ability to comment on language as an entity.
16. Learnability	This feature means that the speaker of one language may acquire another one.

Language and Gender

Sex is a biological category, while gender is a social construct based on sex. Men and women are biologically distinct but gender differences such as dress, behavior, and language are culturally acquired distinctions. What is masculine or feminine in one culture may not be so in another. A fundamental question related to language and gender is whether or not men and women speak differently. In fact, the language of men and women does differ. The way in which it differs falls into several domains. Robin Lakoff's *Language and women's place* (1975) is the groundbreaking study that called attention to differences between men and women's speech habits and the effect that these differences had in terms of power. In the years since its publication, however, several of its claims have been revised based on subsequent empirical research. Various studies show that men and women differ in their conversational styles. Jenny Coates (b. 1942) makes a number of observations about gender differences based on her own research and that of others (Coates, 1993: 138–139). First, context determines speech usage. All-female discourse and all-male discourse is quite different from male–female communication. In male–female conversations, men tend to interrupt women more often. Women employ minimal responses to show support for the speaker and women use hedges (linguistic forms such as *sort of* and *perhaps*). Men, on the other hand, talk more, swear more, and use more commands and imperatives. To be sure, factors such as social status and power affect speech patterns and usage.

Animals and Language

Another issue that merits discussion in this article is whether or not animals other than *Homo sapiens* possess the faculty of language. In his discussion of animal versus human communication, Hockett (1960: 93–95) shows that there are significant differences. A number of psychologists who work with gorillas and chimpanzees argue that these primates can acquire language, in this instance, a form of sign language or artificially created languages. Some of the noteworthy experiments include those listed in Table 2.

Despite claims of communicating with primates via a modified form of human sign language, Sebeok (1981) demonstrates quite persuasively that the animals involved in such experiments are not using language, but rather they are simply responding to subtle cues encoded by the trainer. The most famous claim of interspecies communication is Clever Hans, an Arab stallion from Russia purchased by a man identified as Wilhelm von Osten in 1900. The horse received its name because of a reputed ability to perform mathematical operations as well as an ability to read and to spell. In reality, the horse simply responded to cues from the trainer. The expression 'Clever Hans effect' means that animals behave in certain ways that make it appear to their human captors that they are engaged in language-like functions (Sebeok, 1981: 162–167). Animal–human communication is at best, as Sebeok argues persuasively, conditioned response and not true human linguistic interaction. It should be noted here that Sebeok (1963) coined the neologism 'zoosemiotics' to refer to animal communication. This refers to the sensory signs that animals use to communicate specific messages to other species and to their own species concerning reproduction, territoriality, danger, and so forth.

Semiotics Versus Linguistics

Thomas A. Sebeok (1920–2001) raises an important and sometimes controversial question when he discusses the relationship of semiotics (the study of signs and their uses in representation) and linguistics (the study of language and its usage in culture). Are they coordinate interdisciplines, or is one subordinate to the other? (Sebeok, 1994: 105). Sebeok notes that most would respond that semiotics and linguistics are coequal or that semiotics is subordinate to linguistics. In his carefully crafted response, Sebeok (1994: 106–107) argues that, in fact, linguistics is the subordinate interdiscipline because semiotics deals with the communication of any messages (verbal and nonverbal), while linguistics is concerned only with verbal communication in human languages.

Table 2 Primates and their trainers

Primate	*Trainer*
1. Washoe (chimpanzee)	R. Allen Gardner (b. 1930) and Beatrice T. Gardner (1933–1995)
2. Sarah (chimpanzee)	David Premack (b. 1925)
3. Lana, Sherman, and Austin (chimpanzees)	Duane Rumbaugh (b. 1929) and Emily Sue Savage-Rumbaugh (b. 1946)
4. Koko (gorilla)	Francine Patterson (b. 1947)
5. Nim Chimpsky (chimpanzee)	Herbert S. Terrace (b. 1936)

Linguistic Theory

Linguistic theory has become an important topic for discussion since the 19th century. Because linguistics is considered to be a science, it requires a theoretical apparatus to explain how language functions. It is important to understand the notion of theory. A theory is a hypothesis designed to explain natural phenomena after we observe and describe the basic facts, in this instance, linguistic facts. A credible theory must have the following traits: (1) It must be testable; (2) it must be able to predict; (3) it must be simple; and (4) it must be comprehensive. Though these aspects of theory seem relatively simple, they are difficult to achieve, especially for something as complex as human language.

In the 20th century, linguistic theory received its greatest support from Noam Chomsky (b. 1928), the MIT professor of linguistics. His earliest monograph, *Syntactic structures* (1957) deals with linguistic theory as it relates to syntax. Chomsky envisioned the grammar of a language as a theory of language. Prior to Chomsky's assertion, linguistics was considered to be scientific in the sense of being an accurate description of an observed language. Linguistics is intended to be objective and not prescriptive, i.e., it does not make judgments about how people use their language. Its primary goal is to describe accurately what people say. Chomsky's contribution to linguistic theory goes beyond observation and description to an explanation of language. Chomsky (1965: 4) makes a distinction between competence, the innate ability to produce the sentences of a language and performance, the actual use of language in concrete situations. Chomsky focuses on competence in his theory, while previous linguists focused on performance. He also argues that language is an innate capacity of all humans.

Philosophical Foundations of Linguistics

A review of linguistic research in the 20th century shows that there are two philosophical approaches. First, there is empiricism, the view that we acquire our knowledge of the world through experience, or through the senses. The second is rationalism, the notion that we are born with certain basic knowledge. Throughout the last century, there has been an alternation between these two philosophies. At the beginning of the century, Leonard Bloomfield (1887–1949) published *Introduction to the study of language* (1914), which espoused a rationalistic view of language that Bloomfield called a mentalist viewpoint. Two decades later, under the influence of the behaviorist psychologist, Albert Weiss (1879–1931), Bloomfield completely revised his linguistic perspective to one that he called a mechanistic approach, or an empiricist one in *Language* (1933). This position held sway for nearly a quarter of a century and it finds its most complete articulation in *Verbal behavior* (1957) by the behavioral psychologist, B. F. Skinner (1904–1990). Noam Chomsky faulted every aspect of behaviorist analysis of language, especially its focus on external rather than internal features in his critique of Skinner's book in the journal *Language* (1957, 35:26–58). The debate centered on an empiricist versus a rationalist account of first-language acquisition. Chomsky's approach prevailed because he was able to account for a wide array of linguistic phenomena that previous theories could not explain.

After another quarter century, one of Chomsky's antagonists, George Lakoff (b. 1941) and his colleague Mark Johnson (b. 1949), proposed a novel approach to linguistic theory in *Metaphors we live by* (Lakoff and Johnson, 1980). This empiricist analysis of language research argues that we acquire and use language by interacting with our environment and this interaction is reflected in our language. Its most complete articulation appears in Lakoff's *Fire, women, and dangerous things: what categories reveal about the mind* (Lakoff, 1987), and Ronald W. Langacker's (b. 1942) *Concept, image, and symbol: the cognitive basis of grammar* (Langacker, 2002).

The Origination of Language(s)

In this section, we discuss three ways in which language may arise. The first deals with the species-specific evolutionary development of language. The second addresses contact vernaculars, languages that develop as the result of the contact of two different cultures, e.g., during periods of colonization. The third refers to the creation of languages by individuals to achieve a specific purpose. Finally, we address language death or the loss of a language for various reasons.

Glottogenetics, the science devoted to the study of the origins of language, has come in and out of favor over the years. Because of the plethora of studies on this topic in the 19th century, the Linguistic Society of Paris banned further discussion of it in 1866. Nearly a half century later, in 1911, the Philological Society of London did the same. Nevertheless, in the last quarter of the 20th century, a renaissance of interest in the origin of language took place, as evidenced by the work of Derek Bickerton (b. 1926) *The roots of language* (Bickerton, 1981), and Philip Lieberman (b. 1934) *On the origins of language* (Lieberman, 1975).

Danesi (b. 1946) outlines some of the most common glottogenetic theories of language origination (Danesi, 1993: 5–29). There are at least seven hypotheses:

1. traditional;
2. Otto Jespersen's (1860–1943) enumeration including his own theory;
3. echoic theories;
4. gestural theories;
5. pragmo-communicative theories;
6. scientific approaches;
7. genetics.

Traditional theories about the origination of language include sporadic crude and bizarre experiments designed to determine the first, or earliest human language. These experiments involved the isolation of children from all linguistic contact. Prominent among those who sponsored such cruel trials were Egyptian King Psamtik (663–610 B.C.E.), The Holy Roman Emperor Frederik II (1194–1250), and James IV of Scotland (1473–1513).

Conventional theories of language origins, those that have existed for the longest time, include Otto Jespersen's discussion of four theories (1922: 412–442). First, there is the *bow-wow* hypothesis, which states that speech arose from imitation of natural sounds. Second, the *pooh-pooh* hypothesis, which argues that speech originated from instinctual grunts of pain and pleasure. Third, the *ding-dong* hypothesis, which states speech came about from instinctive oral responses. Fourth, the *ye-he-ho* hypothesis, which asserts that there was a basic need for pre-humans to communicate with their own kind. Jespersen himself offers his own theory, namely, that speech results from the expression of love, play, and affective feelings that developed into song and related forms.

Echoic theories of the origination of speech involve the efforts of pre-humans to imitate the surrounding sounds of these primitive creatures. Certain research into sound symbolism, or onomatopoeia, suggests that there may be shared imitative sounds. Gestural theories argue that the use of gestures, especially the hands, or visual symbolism may account for the origin of language. Pragmo-communicative theories of language origination refer to the development and use of language as necessary to function in society. Scientific approaches to glottogenetics include work by Morris Swadesh (1909–1967) and others who suggested a single source language for the world's languages. Evidence includes scientific principles to reconstruct a proto-language. The comparison of different languages seeks to determine commonalties in the world's languages.

More recently, an evolutionary approach to language origination has become prominent. This approach, represented by Philip Lieberman (1975), argues that language is a mechanism of growth and change that allowed the brain to develop in ways that would enhance survival. Lieberman argues that the development and adaptation of language in the human species are a result of brain structure that allowed humans to generate abstract concepts. The genetic approach to language development is closely associated with Noam Chomsky (b. 1928), who believes that language is a kind of physical organ with grammatical parameters that determine the form of a language.

Pidgins and Creoles

We are able to determine with great accuracy the birth of pidgins and creoles (see McMahon, 1994: 253–283 for a useful discussion). It is necessary to define these terms first. Pidgins are contact languages. They often result from the contact of colonial powers with indigenous populations in their quest to expand their territory. Thus, many pidgins are to be found in Africa, the Caribbean, and Asia. Pidgins result from the contact, most often, of European nations with other nations. In this sense, a pidgin is not the first language of either the conqueror or the conquered. Pidgins represent an effort at some sort of basic communication on the part of both groups. Pidgins are restricted in their usage, e.g., for certain functions such as commerce or trade. Pidgins simplify and restructure two existing languages to create a new language system. The theories about the origin of pidgins are numerous. They include the following basic hypotheses (McMahon, 1994: 255–257):

1. a common nautical language;
2. independent but analogous development;
3. substrate hypothesis, i.e., pidgins combine colonialist vocabulary with indigenous grammar;
4. the monogenetic hypothesis, i.e., all pidgins derive from Sabir a 15th-century Portuguese pidgin;
5. baby-foreigner talk hypothesis, i.e., pidgins represent a simplification of the two languages in contact.

Some linguistic features of pidgins are the following: simple, reduced phonology, reduced vocabulary, no inflectional morphology, and fixed word order.

Creoles, on the other hand, are quite different. First, they represent a highly developed and sophisticated linguistic system with native speakers. Their phonology, morphology, syntax, and lexicon have all of the characteristics of ethnic languages. In this sense, creoles are distinct from pidgins. They still

have a relatively simple phonology but their vocabulary is rich. They have an array of inflectional affixes to indicate plural formation and verb tenses. Furthermore, their syntax is more elaborate and they have embedded sentences and variable word order. Subsequently, they spawn grammars, dictionaries, and creative literature. Some examples of creole languages are the following:

1. Creole Tok Pisin, which means 'speech business' is one of the official languages of Papua, New Guinea and is an admixture of English, Melanesian languages, and German;
2. Haitian Creole, a combination of French and African languages;
3. Hawaiian Creole, a combination of English, Hawaiian, and various Asian languages.

The study of word origins is known as etymology.

Pidgins and creoles are important examples of the origination of new languages that are both accessible to researchers and provide insights into how language first developed. Furthermore, they offer concrete cases of language change in progress. Bickerton, whose work on the origins of language was mentioned above, uses data from pidgins and creoles to support his linguistic bioprogram hypothesis, i.e., all people possess an innate syntax and semantics component.

Planned Languages

Just as we can document the emergence of pidgins and creole languages, we can do the same for created languages, often to the very year of their inception, A planned language is one that is deliberately created by an individual for a specific purpose. At least 900 language schemes have been devised since the 16th century. Umberto Eco (b. 1932) writes about planned languages in *The Search for the perfect language* (Eco, 1995). In that volume, Eco discusses four types of languages:

1. the rediscovery of languages (Hebrew);
2. the reconstruction of languages (Indo-European);
3. languages constructed by individuals including *a priori*, and *a posteriori* languages;
4. so-called magic languages (secret languages endowed with properties of mysticism).

There are essentially two basic types of planned languages. First, there are *a priori* languages, those that are composed entirely of invented elements not found in any existent language. Several examples of this type of language exist. One is Timerio invented in 1921 by a Berlin architect named Tiemer who used a numerical scheme for his *a priori* language. Thus, in Timerio, 1-80-17 means "I love you."

A second general category of invented language is called *a posteriori*. These languages draw upon the grammar, vocabulary, and syntax of natural languages. The best exemplar of this type of planned language is Esperanto, which merits some discussion here because it is the only successful language of its kind. Esperanto is the invention of Dr L. L. Zamenhof (1859–1917), a Polish physician. As a Jew, he suffered at the hands of anti-Semites during his childhood. As a result of his negative experiences, he concluded that a universal auxiliary language would enhance communication and reduce animosity. In fact, his language had two purposes. The first is value-oriented, i.e., contributing to world peace, justice, and the brotherhood of humanity. The other is its norm-orientation, i.e., the use of the language for commerce, travel, and international conferences. The Esperanto language took its name from the pseudonym adopted by Zamenhof to protect his identity. It means "one who hopes." Its vocabulary derives from Romance, Slavic, and Germanic languages. Esperanto has 16 basic rules in its grammar. It uses many prefixes, infixes, and suffixes to create new words from its basic core vocabulary. The language is the only successful created language and it has a rich literature (original works and translations of the most important literary works from ethnic languages). The current number of Esperanto speakers worldwide is estimated to be between 120,000 and 2 million, depending upon the method of calculation. No other planned language has achieved this degree of success.

The following brief selection is from the poem entitled 'La Espero' ('hope') written in 1890 by Zamenhof (Table 3), and considered to the Esperanto anthem. My English translation appears next to it.

Table 3 The Esperanto anthem 'La Espero' '*Hope*'

La Espero	Hope
En la mondon venis nova sento,	*A new feeling came into the world,*
Tra la mondo iras forta voko;	*A strong voice sounds through the world;*
Per flugiloj de facila vento	*On wings of a slight wind*
Nun de loko flugu gi al loko	*May it now fly from place to place.*

Language Death

While glottogenetics is concerned with the origination of language, at the other end of this continuum, the phenomenon of language death has become a matter of growing ecological concern. Wallace F. Chafe (b. 1927) was among the first to

comment on the eventual loss of numerous indigenous languages in North America (Chafe, 1962).

The sociolinguist, Tove Skutnabb-Kangas (b. 1940) addresses the issue of language death in great detail. She (Skutnabb-Kangas, 2000: 46–50) presents a state-of-the-art assessment of the status of endangered world languages today. She cites Michael Krauss (b. 1934) who proposes three categories of languages (Krauss, 1992: 5–7):

1. moribund, which are languages that children no longer acquire;
2. endangered, which are languages that will likely not be learned by children in this century;
3. safe, which do not belong to the first two categories.

David Crystal (b. 1941) poses the question (Crystal 2000: 26–67): "Why should we care about language loss?" He provides five reasons for preserving languages:

1. they express ethnic identity;
2. they preserve history, especially in oral cultures;
3. they contribute to our knowledge of the world;
4. they are interesting expressions of humanity;
5. they contribute to diversity.

Among Crystal's (2000) recommendations for the preservation of endangered languages are the following:

1. increase the prestige of endangered languages in the superordinate community;
2. ensure their presence and use in education;
3. enhance their power in the community;
4. ensure that speakers of endangered languages are literate in that language;
5. use electronic technology for their dissemination.

The reasons for language death are varied. One reason may be a shift from an indigenous language to a recently introduced one, e.g., by a colonialist power, and it makes sense to acquire the new language for practical purposes. In this situation, there is likely to be a period of sustained bilingualism. Second, there may be sudden death. MacMahon (b. 1964) (1994: 286) notes the example of the loss of the indigenous languages Lenca and Cacaopera in El Salvador in 1932 when Indians were murdered and the use of either language represented mortal danger to the survivors. McMahon (1994: 287–291) also discusses language suicide that involves two related languages, one prestigious and the other not. This process is sometimes labeled decreolization. A classic example is the competition of English with Guayanese Creole in the western hemisphere. A third example is language murder. Nancy Dorian (b. 1936) has researched East Sutherland Scottish Gaelic, which is rapidly losing its battle for survival to English (McMahon, 1994: 294–303).

The loss of language means the loss of an entire culture, or at least the loss of that culture as an ongoing source of creativity and inventiveness though cultural vestiges such as literature and art will survive. Language death occurs for a variety of reasons. It may be a deliberate act of a national government in order to promote unity and solidarity and to exterminate dissidents of a minority group who may seek autonomy or independence, e.g., Francisco Franco's (1892–1975) suppression of regional languages in Spain (1939–1975).

Language Change

We are all aware of language change when we travel. The English of Boston, London, Melbourne, and Calcutta are quite distinct. When we take survey courses in English literature, we see different forms of English in Beowulf, Romeo and Juliet, and a modern novel.

Language is not static. It changes from place to place and generation to generation. Spatial language change involves synchronic linguistics, i.e., changes in contemporary language. Historical change is diachronic change, i.e., language changes over time. Ferdinand de Saussure (1957–1913) added to our knowledge of language change by establishing the distinction between diachronic and synchronic language analysis.

There have been many efforts to explain language change. In the 19th century, German linguists proposed linguistic laws to explain diachronic language change. Rasmus Rask (1787–1832), for example, first proposed a law of sound correspondences in Indo-European. Subsequently known as Grimm's law after Jakob Grimm (1785–1863), this set of sound correspondences specifies how certain consonants appeared in modern Germanic languages such as English, Dutch, and Swedish. Although Grimm's law explained many changes, it had exceptions. The Danish linguist, Karl Verner (1846–1896) was able to find the generalization to Grimm's law to make it exceptionless. This approach to language change was labeled comparative grammar by the German scholar Friedrich Schlegel (1772–1829).

André Martinet (1909–1999) suggested a tension between the speaker and the hearer as a factor in language change, i.e., the speaker wants to expend minimal effort in articulation while the hearer seeks maximum differentiation. In *The sound pattern of English* (Chomsky and Halle, 1968), Noam Chomsky and Morris Halle (b. 1923) argue that language

change is grammar change, i.e., there are changes in the grammar of a language that account for phonological change. These grammatical changes may involve rule addition. These authors provide evidence of this type of change in English, especially the Great English Vowel Shift occurred because of rule addition.

In the past four decades, sociolinguists, those who study the use of language in various social contexts, have added to our knowledge of language change. The approach of the sociolinguist is to look at socioeconomic factors to try to explain language change. Factors that come into play include prestige of a language register, educational background, group identity, interethnic contact and influence, to name but a few (McMahon, 1994: 225–252).

Pragmatics

Levinson (b. 1947) points out that the modern usage of the term 'pragmatics' derives from Charles Morris' (1903–1979) work (Morris, 1938) on semiotics (Levinson, 1983: 1). Levinson (1983: 5) goes on to note that it is difficult to provide a succinct yet comprehensive definition of the term 'pragmatics.' Levinson (1983: 9, emphasis in original) provides the following definition of pragmatics that is a useful one: "Pragmatics is the study of those relations between language and context that are **grammaticalized**, or encoded in the structure of a language."

In simple terms, pragmatics is the subdomain of linguistics that examines how language is used in specific social contexts. Our parents are usually responsible for teaching us pragmatics and they often do so when they direct us to use certain words and phraseology in particular situations, e.g., when we should address people by their first name, or by their last name with an appropriate title (Mr, Mrs, Miss, Ms, Doctor, Reverend, and so forth). This particular form of pragmatics is labeled politeness. Politeness is often encoded in languages of the world. When you study the Romance languages, for example, you learn that there two ways of expressing the English word 'you' (English had two ways also 'you' and 'thou,' which still exist in Amish English). In French, there are the forms '*tu*' and '*vous*'; in Spanish there are the forms '*tú*' and '*usted*'; and in Italian '*tu*' and '*Lei*.' These familiar and polite forms use different verb forms and this is a very difficult concept for Anglophones to acquire, especially when they are acquiring a second language.

When we want someone to do something for us, we use different forms of language to achieve our objectives. There is a hierarchy of linguistic forms that represents the social or official status of the speaker and hearer. These formal differences correspond to a scale politeness (Brown and Levinson, 1987) that conforms to social power structures. Thus, there is a direct correlation between linguistic form and power status. The scale of directness moves from most direct to most subtle with a series of intermediate points along this scale. First, there are direct commands or imperatives that coincide with a superordinate–subordinate status. Examples of this linguistic context would include police and military personnel who use the familiar imperative to make the hearer comply with a certain command, e.g., *stop*! or *halt*! Second, are performatives, e.g., *I'm asking you to move that chair.* Third, there are obligation statements, e.g., *you'll have to move that chair.* Fourth, there are suggestions, e.g., *why not move that chair.* Fifth, there are hints, e.g., *the chair is still there*. Finally, there are mild hints, e.g., *we don't want too much furniture in this room*. Table 4 illustrates such a scale in which the first is most direct and the last is least direct.

Imperatives, which are the most direct requests, have a certain linguistic form, i.e., a verb with no subject, e.g., *move!* Milder requests take the form of statements, e.g., *there is more space in the living room* (request that someone move). Likewise, questions may also be requests, e.g., *where's the salt?* (meaning, 'please pass the salt'). In *Toward a linguistic theory of speech acts* (Sadock, 1974), Jerrold Sadock (b. 1942) calls the latter whimperatives because they frequently uses a *wh*-question (what?, where?, why?, and so forth) word in their formulation.

Jenny Thomas (b. 1948) points out that when we acquire a second language, we often fail to learn certain pragmatic strategies such as the appropriate forms of politeness in the target language that she calls pragmatic failure (1983: 99). This means that although we may have total command of the phonology (the study of the sound system of a

Table 4 Scale of politeness for requests

Most direct	→	→	→	→	*Least direct*
Direct command (imperative) 1	Performative statement 2	Obligation statement 3	Suggestion 4	Hint 5	Mild Hint 6

language) and syntax (the study of the order and arrangement of the words of a language) of the second language, we may not know which forms to use in the appropriate contexts, with the result that we project an image of an abrupt or impolite person. Even the most fluent person may be unsuccessful if he or she ignores these subtle pragmatic rules of politeness.

Neurolinguistics

Neurolinguistics is the study of the relationship of the brain to language. The brain orchestrates the complex operations involved in comprehending and producing language. When it suffers damage (brain attacks or strokes, Alzheimer's disease, and other maladies), it fails to direct orderly speech operations. This results in failed comprehension by the speaker and it also means that the hearer will not be able to understand the messages sent by a speaker with a neurolinguistic affliction.

As we noted previously, physicians are semioticians. Their task in diagnosis is to decode what outward symptoms stand for. A physical symptom signifies something other than itself. In medicine, this semiotic tradition goes back to the ancient Greeks. Hippocrates (460–377 B.C.E.) provided the first definition of a symptom, and Galen (139–199 C.E.) followed this tradition. Physicians examine superficial signs to arrive at a diagnosis. In fact, the classical Latin symptoms of disease included the following:

1. *rubor* 'redness';
2. *dolor* 'pain';
3. *fervor* 'fever';
4. *tumor* 'swelling.'

These symptoms may correspond to a variety of maladies including allergies, infections, and broken bones, to name but a few possibilities. Symptomatology remains an important part of the medical curriculum today. It exemplifies semiotics perfectly because the symptom stands for something else. A precise determination of what symptoms stand for is a physician's daily task.

Speech pathology is that branch of medicine that seeks to provide clinicians with an exact set of correspondences of symptoms or signs to specific physical malady. In *Linguistic encounters with language handicap* (Crystal, 1984), David Crystal discusses the need for more work on symptomatology in the domain of speech pathology. In particular, Crystal notes the need for good transcription of speech dysfunction and standardized typologies of these maladies.

Aphasia (Greek 'no speech'), is partial or total inability to produce meaningful speech. It is occasioned by damage to specific areas of the brain. Specific parts of the brain control speech functions. One is Broca's area, which controls the musculature of thethroat and mouth for speaking. This area was named for the French scientist Pierre-Paul Broca (1824–1880) who found a correlation between aphasia and damage to a specific area of the brain, reported in a paper published in 1861. Broca's aphasia has several manifestations: limited word use, frequent pauses, short sentences, and nonfluent speech. In 1874, 13 years later, the German scientist, Carl Wernicke (1848–1904) published a paper that linked a specific type of speech disorder to damage in the area now called Wernicke's area. Wernicke's area controls speech comprehension. In simple terms, its manifestations include the loss of comprehension of spoken language, the inability to write, the loss of silent reading comprehension, and the distortion of speech.

When people lose the ability to produce or comprehend speech, they are unable to communicate appropriately with their family, friends, and associates. Such speech disfluencies affect social interactions and they need to be addressed clinically.

Language Games

Language games constitute an important part of human behavior. Children engage in them and they are a meaningful part of first-language acquisition, while adults enjoy them for their entertainment and creative value. Language puzzles are enormously popular and they may be found in magazines dedicated to this activity such as *Games* or as regular features in serious journals such as *Scientific American*. Puzzles may be nonverbal or verbal. Mathematical problems and logical problems are nonverbal though they may be presented in story form to make them easier to comprehend verbally. Our interest here is verbal puzzles, games that you can play with language. We provide several examples of such games here (see Danesi, 2002 for further discussion). Language games provide amusement and develop linguistic skills for those who create them or play them. An excellent language exercise is to try to create your own language games.

Riddles

A riddle is a statement or a question intended to puzzle the hearer. The classic riddle is the Riddle of the Sphinx, believed to be the oldest known puzzle. The Sphinx was an ancient mythological monster.

This beast lurked near the entrance of the city of Thebes and, under penalty of death, required of those who wished to enter that they answer the following riddle: "What has four feet in the morning, two at noon, and three at twilight?" The answer, of course, is man, because he crawls on all fours as a child, walks upright as an adult, and uses a cane in later life. Books of riddles abound and riddling is a favorite activity of children.

Palindromes

Palindromes are phrases or sentences that may be read the same forward and backward. They have a long history, and the originator of this form of linguistic amusement is reputed to be Sotades, a poet from Alexandria in the 3rd century B.C.E. The classic example of this type of word game is "A man, a plan, a canal – Panama."

Anagrams

An anagram (Greek for reversed letters) is a language game in which the letters of an existing word are rearranged to create an entirely new word (spit → tips) or a rearrangement of the letters of an existing word or phrase to create a new phrase or sentence. With respect to the second type of anagram, it is common to use the name of a well-known figure to create a phrase or sentence that captures the essence of that person. The rearrangement of the title of a well-known piece of literature may encapsulate its content. Two examples of this type of language game come from Bergeson's (1973: 77, 81) book *Palindromes and anagrams*: the title of Jules Verne's (1828–1905) famous novel *Twenty thousand leagues under the sea* → "Huge water tale stuns. End had you tense" and William Shakespeare (1564–1616) → "I ask me, has Will a peer?"

Word Ladders

Word ladders, sometimes called *doublets* or *word ping-pong*, involve the change of a single letter of a word by changing one letter to create an entirely new word. Figure 1 is a simple example of this type of language game (each letter change appears in boldface type).

STILL
ST**A**LL
STAL**K**
STAL**E**
STA**K**E
S**N**AKE

Figure 1 Word ladder (word doublet, word ping-pong).

Word Squares

Word squares are words of equal length that read the same vertically and horizontally. The longer the word, the more challenging is the task. Figure 2 is a typical example of this type of language game.

MET
EKE
TEN

Figure 2 Word square.

Crossword Puzzles

Arthur Wynne (1862–1945), a journalist for the *New York World*, is credited with the invention of the now ubiquitous crossword puzzle, although he called it a word cross. His first puzzle appeared in that paper on December 21, 1913. The best known enigmatologist is Will Shortz (b. 1952), successor to Eugene T. Maleska (1916–1993), who is responsible for the prestigious *New York Times* crossword puzzle.

The crossword puzzle consists of a set of clues that may be thematic and that can be quite challenging. In fact, there is an annual crossword puzzle contest. The reader must write the appropriate word either vertically or horizontally. The simple crossword puzzle in Figure 3 illustrates this type of word game. The clues relate to linguistics and language and the answers can be found in this article.

Lipograms and Univocalics

A lipogram is a language game in which the author writes a creative work by deliberately omitting a particular letter (lipogram) of the alphabet or by using a single vowel (univocalics). The Greek poet Lasus (538 B.C.E.-?) is perhaps the first author to write in this fashion. He wrote two poems, an ode to the Centaurs (Greek mythological animal with head and arms of a man and the trunk of a horse) and another work to the Ceres (goddess of agriculture) in which the letter *s* was deliberately omitted. Many other writers have imitated this practice.

Oulipo, alternatively written as *OuLiPo* (abbreviation of the French *Ouvroir de Littérature Potentielle* 'work of potential literature'), is a group of creative writers in France who seek to write distinctive creative works subject to various constraints and regulations (Motte, 1998). The origin of this group may be traced to a meeting of diverse professionals who convened at Cerisy-la-Salle for a colloquium on the work of the French author Raymond Queneau (1903–1976) in September 1960. One intriguing example of their literary tradition is the novel *La disparition* (Perec, 1969) written by George Perec (1936–1982). What makes this literary work so interesting is the

fact that the author wrote the novel (300 pages long) without using the vowel *e*. Furthermore, Gilbert Adair (b. 1944) translated the original French work into English as *A void* (Adair, 1995), again without using the letter *e*. In his review of the English translation in *Time* magazine (Gray, 1995: 74), Paul Gray (b. 1930) did not use the letter *e*. The following passage of Gilbert Adair's English translation of Perec's novel (Adair, 1995: 3) illustrates this technique.

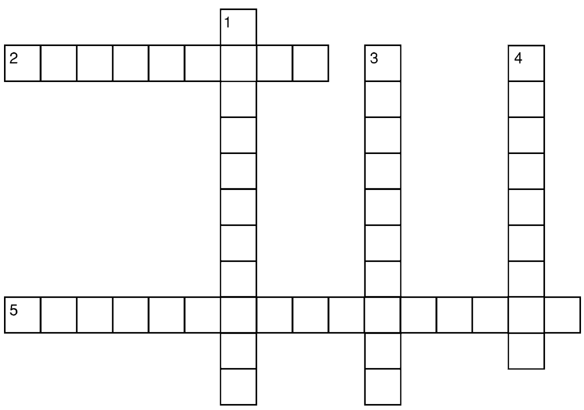
> Incurably insomniac, Anton Vowl turns on a light. According to his watch it's only 12.20. With a loud and languorous sigh Vowl sits up, stuffs a pillow at his back, draws his quilt up around his chin, picks up his whodunit and idly scans a paragraph or two. . . .

Names

The study of names is called onomastics. Names are one of the most significant aspects of language. In order to refer to people, objects, concepts, and places, they must have a name. Names pervade our linguistic existence. We use names to label and categorize the world around us.

Personal Names

Personal names include family names, or surnames, and given names (first names and middle names). Names give us identity and it is an important ritual worldwide. Surnames denote lineage, while given names provide us with personality and individuality.



Across

2 Study of signs
5 Study of language and the brain

Down

1 Study of language
3 Study of names
4 Study of meaning

Figure 3 Linguistic crossword puzzle.

Given names often reflect name fads, i.e., certain names are popular during certain eras and may fall out of favor during others. The popularity of certain names often reflects those of well-known personalities of a particular time, most often entertainers (film stars, singers, and so forth). Parents often seek to give their children names that do not deviate from perceived social norms of a particular era so that their children are not burdened by unpleasant teasing from their peers. The U.S. Social Security Administration provides a comprehensive listing of given names that dates to 1880. We have provided a list of the most popular male and female given names for both 1900 and 2003 in Table 5. Although these names represent a single country, similar lists may be found for other nations as well.

Name Changes

Name changes represent another aspect of onomastics. The reasons for changing one's name are varied. People may not like their given name, especially if it is unpopular or somehow perceived to be aesthetically displeasing. A family name may have an unpleasant connotation, e.g., Hogg, Hitler, Stalin, and so forth. Entertainers may have names that are difficult to pronounce or spell. Furthermore, they seek a name that is easy to remember and that has some positive connotation, e.g., Issur Danielovich Demsky (Kirk Douglas [b. 1916]), or Francis Gumm (Judy Garland [1922–1969]). A third reason for such changes may be for criminal ends, e.g., to defraud someone of money, or for the purpose of avoiding identification and detection. A name change is a significant semiotic act because it alters our personal and public identity.

Language of Oppression

The language of oppression involves the naming and labeling of groups in distorted and negative ways so that the dominant and oppressing group may have control over the oppressed one. Haig Bosmajian (b. 1929) notes that the power to name one's foe means that you can define it (Bosmajian, 1974: 5–6). This linguistic power transfers self-definition from the oppressed group to the oppressor, often with severe consequences.

Language may be used to express prejudice and discrimination. It may encode prejudice (negative attitudes based on erroneous information about a particular group), or discrimination (acts intended to cause harm to a group, e.g., failure to hire, or even to eliminate an entire group of people as was the case with The Holocaust).

Table 5 Most popular male and female given names in the United States

Boys' names (1900)	*Girls' names (1900)*	*Boys' names (2003)*	*Girls' names (2003)*
1. John	1. Mary	1. Jacob	1. Emily
2. William	2. Helen	2. Michael	2. Emma
3. James	3. Anna	3. Joshua	3. Madison
4. George	4. Margaret	4. Matthew	4. Hannah
5. Charles	5. Ruth	5. Andrew	5. Olivia
6. Joseph	6. Elizabeth	6. Joseph	6. Abigail
7. Frank	7. Marie	7. Ethan	7. Alexis
8. Henry	8. Rose	8. Daniel	8. Ashley
9. Robert	9. Florence	9. Christopher	9. Elizabeth
10. Harry	10. Bertha	10. Anthony	10. Samantha

Frequently, words to describe these acts become part of the language long after the behavior associated with them have occurred. In the past 40 years, at least two terms that refer to categories of long-standing prejudice and discrimination, sexism, and ageism, have entered the English language and now appear as dictionary entries.

The language of American Indian derision, and the language to disparage the disabled are other domains in the language of oppression. The derisive language aimed at minority groups often reflects the behavior or the appearance of the targeted group. The linguistic terms to deride these groups, often metonymic (the use of the part for the whole) in nature, serve to focus on some superficial aspect the oppressed group's appearance such as skin color. These metonyms may also allude to cultural habits such as clothing, eating habits, or some other equally unimportant external custom of the group.

Bosmajian (1974) notes that legal language often encodes prejudice and discrimination. In fact, officially promulgated laws provide strong evidence for the language of oppression, as evidenced by anti-Semitic legislation and official decrees in Nazi Germany, sexist language in U.S. laws, and the racist Jim Crow laws of the post Civil War in the United States, and U.S. court cases that excluded Native Americans from civil rights in the United States.

The Language of Advertising

Danesi (2000: 6) defines advertising as "[t]he making of public announcements designed to promote the sale of specific commodities or services." Advertising has a long history, and it can be traced to ancient times. This form of commercial publicity was at first oral. Prior to the invention of the printing press, people proclaimed their wares in public by shouting. Once relatively inexpensive paper reproduction became available, printed advertising became ubiquitous. Advertising as a formal corporate service first appeared in the 19th century. Volney P. Palmer (1799–1864) started the first agency in 1841. At the end of the 19th century, advertising was a full-time business. At that time, branding of products was the most important task. This changed in the 20th century when companies sought not only to identify their products but also to persuade the public to buy their products and services. Danesi (2000: 7) points out that advertising is now "a dominant form of social discourse."

Advertising thus involves the use of language to advocate, promote, persuade, engage, and interact with potential customers. Advertisers accomplish this task with language as well as other means. We shall examine this process now. Prior to discussing language *per se*, it is necessary to consider how advertisers achieve their goal of convincing the public to pay attention to their product, and then, to buy it. We receive information about our external world through our five senses (hearing, sight, smell, taste, and touch). In its earliest form, much of advertising involved a combination of textual and visual materials. In the 20th century, this process became more sophisticated. Advertisements for perfumes and colognes included magazine inserts that allowed people to smell the product. Likewise, certain popup advertisements provided a three-dimensional component to advertising. It has become impossible to escape advertising in the 21st century. It is literally everywhere. In the print media, it may be found in the traditional places such as magazines, newspapers, flyers, and so forth. In the broadcast media, it appears in television, radio, and film. Product placement, the subtle or not so subtle insertion of certain well-known products in films and television programming, means that advertising is ubiquitous.

Roland Barthes (1915–1980) specifies two types of messages encoded in advertising, the quintessential semiotic text (Barthes, 1964). These are the linguistic message and the visual image. The latter involves the

visual depiction of the product or service and a second, implied or unstated connotation, i.e., what the product means to the consumer. Tobacco advertising has been quite effective in delivering this hidden, coded message. Forty years ago, the Marlboro man depicted a virile cowboy who personified masculine virility. By association, using this product would bestow on the user these same qualities. More recently, tobacco products have evoked a notion of coolness (the image of a person who is popular and who engages in peer-approved activities) to effect the connotative message. In fact, certain messages involving a hip-hop lifestyle are directed at inner-city youth in an effort to recruit young adults to use these products.

Our interest here is the use of language to promote products and services. As noted above, advertising encodes overt and covert messages to the consumer. At the superficial level, advertising tells the public what the product or service will do: clean, provide easily consumable food, and so forth. At the covert level, advertising evokes popular myths. If we use a product, it will make us resemble, or perhaps become, certain archetypal figures who signify virility, femininity, and so forth. The models and the actors convey this image to the consumer. The verbal messages (including warnings) tell us one thing, while the visual messages implicit messages convey other more significant ones about lifestyle.

Truth in Advertising

Representing a product or service is now subject to some minimal requirements of truthfulness. Most countries have governmental agencies that exert some minimum standards of honesty for product and service promotion. While overt lying is thus not permitted, compliance with legal regulations often involves the skillful use of language and visual accommodation to meet legal requirements. We now know that certain products may be harmful. Smoking cigarettes, for example, may cause death from a variety of diseases, including cancer. When first required to place warning labels on their product, cigarette companies simply placed somewhat ambiguous statements on the packages. Subsequently, these warnings became more specific and prominent as legislation became more demanding.

The warnings and disclaimers in advertising are often hidden through various media artifacts. In print media, they appear in small print at the bottom of a colorful and appealing pictorial presentation. In non-print media, they appear as white, verbal text against a light background for a moment at the end of announcement or in a rapid hyper-speech enumeration at the end of the main message, thus making it difficult to comprehend.

Producers of alcoholic beverages now seek to avoid lawsuits by inserting language about responsible drinking unobtrusively into their advertisements. Advertisements for pharmaceutical products do the same by mentioning sometimes long lists of side effects (sometimes including death) in their product announcements.

Product Names

A trade name is very important. Adrian Room (b. 1933) specifies six criteria for product names (Room, 1982: 15–16):

1. visual comprehension (easy to read);
2. easy to pronounce in most languages;
3. avoidance of unacceptable or absurd meanings;
4. positive references;
5. ease of recall;
6. conformity to legal requirements for official registration.

Trade names fall into three basic groups. The first is the personal name, i.e., the name of the founder of a company, e.g., *Sears*®. The second is a name that represents some special feature of the product, e.g., *Sunkist*® or *Safeway*®. The last is the arbitrary name, i.e., one with no obvious reference, e.g., *Exxon*®.

Other factors go into trade names. Spelling, for example, is important. Certain consonants are preferred, e.g., X (*Xerox*®) and K (*Kodak*®). X is a letter that corresponds to several sounds [z] and [ks]. Other letters, L, M, N, have syllabic value and thus they convey more than one sound. Morphology, of the formation of words, is an important part of product designation. Pharmaceutical products tend to end in suffixes such as *-ine* (*Listerine*®, *Dramamine*®) or *-in* (aspirin). The semantics, or meaning of a product or service name, is also quite important. A product name may involve snob appeal, e.g., *Lord and Taylor*®. Trade names may identify the function of the product or service, e.g., *Chem-Lawn*®.

Car names encode hierarchies that usually correspond to price ranges and status, e.g., *Chevrolet*®, *Pontiac*®, *Buick*®, and *Cadillac*®. Within a given car line, there exist hierarchies also (make, line, model). *Chevrolet*® thus has various subdivisions including *Aveo*®, *Cavalier*®, *Malibu*®, and *Impala*® within its standard line. In his classic sociological class, *The status seekers* (Packard, 1959: 313), Vance Packard (1914–1996) noted that "the automobile makers built up a status hierarchy for automobiles that was well understood by the public."

Concluding Remarks

Language is an evolutionary mechanism that *Homo sapiens* developed to enhance survival. It is a multifunctional tool that allows us to interact as social beings with other people in our daily lives. Various technologies have allowed us to encode language in order to record past events and to refer to previous events and the circumstances that surround them. Written language thus preserves our past. Modern technology also allows virtually instantaneous contact with other people globally.

What is important about language is the fact that it encodes cultural content in various, often subtle ways. Our choice of words and specific grammatical forms represent our perceptions of the world. Our linguistic interactions and usage reveal us to the world. Most of Erving Goffman's (1922–1982) research deals with communication and language, either directly or indirectly (*Behavior in public places*, 1963). He makes the point that all of our social intercourse revolves around language, or even occasionally, on the lack of it, i.e., silence. When we use language, we encode significant amounts of collateral information about ourselves and our beliefs in indirect ways. Likewise, when we decode linguistic messages, we interpret a great deal about the message through word choice and grammatical form. The form of the message sometimes provides more information than its content.

See also: Figurative Language: Semiotics; Language Change and Cultural Change; Pidgins and Creoles: Overview; Pragmatics: Overview.

Bibliography

Adair G (tr.) (1994). *A void.* London: Harvill.

Barthes R (1964). 'Rhetoric and image.' In Barthes R (ed.) *Image–music–text.* New York: Hill and Wang. 32–51.

Bergerson H W (1973). *Palindromes and anagrams.* New York: Dover Publications.

Bickerton D (1981). *The roots of language.* Ann Arbor, MI: Karoma.

Bloomfield L (1914). *Introduction to the study of language.* New York: Henry Holt.

Bloomfield L (1933). *Language.* New York: Holt.

Bloomfield L (1957). *Verbal behavior.* New York: Appleton-Century-Crofts.

Bosmajian H (1974). *The language of oppression.* Washington, DC: Public Affairs Press.

Bourdieu P (1989). *Language and symbolic power.* Raymond R & Adamson M (trs.). Cambridge: Harvard University Press.

Brown P & Levinson S C (1987). *Politeness. Some universals in language usage.* Cambridge: Cambridge University Press.

Chafe W F (1962). 'Estimates regarding the present speakers of North American Indian languages.' *International Journal of Anthropological Linguistics 28,* 162–174.

Chomsky N (1965). *Aspects of the theory of syntax.* Cambridge: MIT Press.

Chomsky N & Halle M (1968). *The sound pattern of English.* New York: Harper and Row.

Coates J (1993). *Women, men and language* (2nd edn.). London: Longman.

Crystal D (1984). *Linguistic encounters with language handicap.* Oxford: Blackwell.

Crystal D (2000). *Language death.* Cambridge: Cambridge University Press.

Danesi M (1993). *Vico, metaphor, and the origin of language.* Bloomington, IN: Indiana University Press.

Danesi M (2000). *Encyclopedic dictionary of semiotics, media, and communications.* Toronto: University of Toronto Press.

Danesi M (2002). *The puzzle instinct: the meaning of puzzles in human life.* Bloomington, IN: Indiana University Press. (This is an excellent overview of language games and the theory that underlies them).

Eco U (1995). *The search for the perfect language.* Fentress J (tr.). Oxford: Blackwell.

Gray P (1995). *Time 145,* 74. (February 6).

Hockett C F (1960). 'The origin of speech.' *Scientific American 203(3),* 88–96.

Hockett C F (1963). 'The problem of universals in language.' In Greenberg J H (ed.) *Universals of language.* Cambridge: MIT Press. 1–29.

Jespersen O (1922). *Language: its nature, development and origin.* London: Allen and Unwin.

Krauss M (1992). 'The world's languages in crisis.' *Language 68,* 4–10.

Lakoff G (1987). *Fire, women, and dangerous things: what categories reveal about the mind.* Chicago: University of Chicago Press.

Lakoff G & Johnson M (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Lakoff R (1975). *Language and women's place.* New York: HarperColophon.

Langacker R W (2002). *Concept, image, and symbol: the cognitive basis of grammar.* Berlin: Mouton de Gruyter.

Levinson S C (1983). *Pragmatics.* Cambridge: Cambridge University Press.

Lieberman P (1975). *On the origins of language.* New York: Macmillan.

McMahon A M S (1994). *Understanding language change.* Cambridge: Cambridge University Press. (Excellent source for pidgins, creoles, and language loss).

Morris C W (1938). 'Foundations of the theory of signs.' In Neurath H, Carnap R & Morris C (eds.) *International encyclopedia of unified science.* Chicago: University of Chicago Press. 77–138.

Motte W F Jr (1998). *Oulipo a primer of potential literature.* Normal, IL: Dalkey Archive Press.

Packard V (1959). *The status seekers.* New York: D. McKay.

Perec G (1969). *La disparition.* Paris: Denoël.

Perec G (1995). *A void.* Adair G (tr.). New York: Harper-Collins.
Room A (1982). *Dictionary of trade name origins.* London: Routledge and Kegan Paul. (This is a useful resource for product and service names in advertising).
Sadock J (1974). *Toward a linguistic theory of speech acts.* New York: Academic Press.
Saussure F (1986 [1916]). *Course in general linguistics.* Chicago: Open Court Press.
Sebeok T A (1963). 'Communicating among social bees; porpoises and sonar; man and dolphin.' *Language 39,* 338–466.
Sebeok T A (1981). 'Smart simians: the self-fulfilling prophecy and kindred methodological pitfalls.' In Sebeok T A (ed.) *The play of musement.* Bloomington, IN: Indiana University Press. 134–209.
Sebeok T A (1994). *Signs: an introduction to semiotics.* Toronto: University of Toronto Press.
Skutnabb-Kangas T (2000). *Linguistic genocide in education or worldwide diversity and human rights.* Mahwah, NJ: Lawrence Earlbaum Associates.
Thomas J (1983). 'Cross-cultural pragmatic failure.' *Applied Linguistics 4,* 91–112.

Languages for Specific Purposes

J Engberg, Aarhus School of Business, Aarhus, Denmark

Defining the Object

The object of study for the branch of applied linguistics that is covered by the heading 'Languages for Specific Purposes' is multi-faceted. The many facets are a consequence of the fact that this discipline looks at (in principle all) aspects of actual communication in specialized discursive domains. To say that discursive domains are specialized means that it is possible to become a specialist in the domain, through education, training, or experience. This definition covers primarily professional areas, but also nonprofessional areas like hobbies. The fact that not a focused part of the human linguistic competence, but the totality of aspects involved in communication in the described settings is the potential object of study, gives studies into the field a very wide range. However, two main poles may be isolated in the landscape of LSP research, around and between which different approaches and projects are located:

- One pole sees LSP as supplementary language skills that are applied when producing texts in specialized situational settings in order to realize specific purposes. The basic assumption is that language is generally used in the texts according to the ordinary rules known to all language users (Language for General Purposes, LGP), but the ways words are connected and the words are specialized. This approach is generally meant by the English term used to designate the object of study, viz., 'Language for Specific Purposes.' The concept centrally behind this term is closely connected to language teaching for professional purposes, where speakers of, e.g., English as a foreign or second language have to learn (additionally) how to use language in areas where they are going to work. In other words, scholars tending to this pole concentrate upon the specific linguistic differences between language used in different settings.
- The other pole is more closely connected to the concept of specialized meaning. The center of this pole is the concept meant by the German term *Fachsprache*, i.e., domain or subject specific language. This concept lays more weight upon the specialized meanings constituting a domain and upon the relations between these meanings and the linguistic choices conventionally made by the agents of the domain. In other words, the view of the scholars tending to this pole is more global, they are rather looking at the language of a domain as a general reflex of the domain and the specialized meanings constituting the domain. This latter view is more discourse oriented in its approach to its object.

The two poles describe two original, fairly different approaches to this object: the first mentioned approach, concentrating upon text production in specialized situational settings, is connected to language teaching, especially to teaching a foreign or second language in connection with vocational training, university studies, etc. A major interest in this approach is to create knowledge about the specific needs to be covered in such specialized language classrooms, in order to make this kind of language teaching as efficient as possible. The study of domain specific language use comes in handy as a linguistic basis for such specialized courses. This kind of language teaching is mostly connected to learning English. For one thing, this occurrence is due to the fact that a number of countries with multiple national

Perec G (1995). *A void.* Adair G (tr.). New York: HarperCollins.
Room A (1982). *Dictionary of trade name origins.* London: Routledge and Kegan Paul. (This is a useful resource for product and service names in advertising).
Sadock J (1974). *Toward a linguistic theory of speech acts.* New York: Academic Press.
Saussure F (1986 [1916]). *Course in general linguistics.* Chicago: Open Court Press.
Sebeok T A (1963). 'Communicating among social bees; porpoises and sonar; man and dolphin.' *Language 39,* 338–466.
Sebeok T A (1981). 'Smart simians: the self-fulfilling prophecy and kindred methodological pitfalls.' In Sebeok T A (ed.) *The play of musement.* Bloomington, IN: Indiana University Press. 134–209.
Sebeok T A (1994). *Signs: an introduction to semiotics.* Toronto: University of Toronto Press.
Skutnabb-Kangas T (2000). *Linguistic genocide in education or worldwide diversity and human rights.* Mahwah, NJ: Lawrence Earlbaum Associates.
Thomas J (1983). 'Cross-cultural pragmatic failure.' *Applied Linguistics 4,* 91–112.

Languages for Specific Purposes

J Engberg, Aarhus School of Business, Aarhus, Denmark

Defining the Object

The object of study for the branch of applied linguistics that is covered by the heading 'Languages for Specific Purposes' is multi-faceted. The many facets are a consequence of the fact that this discipline looks at (in principle all) aspects of actual communication in specialized discursive domains. To say that discursive domains are specialized means that it is possible to become a specialist in the domain, through education, training, or experience. This definition covers primarily professional areas, but also nonprofessional areas like hobbies. The fact that not a focused part of the human linguistic competence, but the totality of aspects involved in communication in the described settings is the potential object of study, gives studies into the field a very wide range. However, two main poles may be isolated in the landscape of LSP research, around and between which different approaches and projects are located:

- One pole sees LSP as supplementary language skills that are applied when producing texts in specialized situational settings in order to realize specific purposes. The basic assumption is that language is generally used in the texts according to the ordinary rules known to all language users (Language for General Purposes, LGP), but the ways words are connected and the words are specialized. This approach is generally meant by the English term used to designate the object of study, viz., 'Language for Specific Purposes.' The concept centrally behind this term is closely connected to language teaching for professional purposes, where speakers of, e.g., English as a foreign or second language have to learn (additionally) how to use language in areas where they are going to work. In other words, scholars tending to this pole concentrate upon the specific linguistic differences between language used in different settings.
- The other pole is more closely connected to the concept of specialized meaning. The center of this pole is the concept meant by the German term *Fachsprache*, i.e., domain or subject specific language. This concept lays more weight upon the specialized meanings constituting a domain and upon the relations between these meanings and the linguistic choices conventionally made by the agents of the domain. In other words, the view of the scholars tending to this pole is more global, they are rather looking at the language of a domain as a general reflex of the domain and the specialized meanings constituting the domain. This latter view is more discourse oriented in its approach to its object.

The two poles describe two original, fairly different approaches to this object: the first mentioned approach, concentrating upon text production in specialized situational settings, is connected to language teaching, especially to teaching a foreign or second language in connection with vocational training, university studies, etc. A major interest in this approach is to create knowledge about the specific needs to be covered in such specialized language classrooms, in order to make this kind of language teaching as efficient as possible. The study of domain specific language use comes in handy as a linguistic basis for such specialized courses. This kind of language teaching is mostly connected to learning English. For one thing, this occurrence is due to the fact that a number of countries with multiple national

languages (often former British colonies like India or Malaysia) use or have until recently used English as a common language, for example, in administrative and judicial settings. Here, there is a strong need for specialized teaching. Secondly, the position English has especially in the scientific world as *the* lingua franca of international contacts and publications means that there is a strong everyday need among students and scholars from all over the world to acquire specific and specialized English language skills. Consequently, this approach has been followed mainly by scholars working in the field of teaching English to adults with specialized professional or academic needs. Due to the distribution of English in the world described above, the approach is not limited to Britain or the United States, but is widely spread in all parts of the world, although predominantly connected to the teaching of English.

The second mentioned approach, the one focusing more on the specialized meanings, has a different root. It grew out of a general interest in sociology and dialectology and was an extension of former studies of different population groups' ways of living through studies of their discourse. Thus, the original interest was wider in its scope. This study lead to an early interest in global models of communication in specialized settings, combining text internal and text external factors in the descriptions and focusing upon global explanations including and combining discoursal features, relevant knowledge, and social backgrounds rather than just finding specialized elements of the discourse. Germany and Austria were the two countries where this kind of study of specialized discourse was first developed, but also the Scandinavian countries have a tradition for investigating specialized discourse from this perspective.

It is important to say that modern LSP research is normally not placed at any one of the two poles, but rather at some place on the continuum between them. And furthermore there is no 1:1 relation between countries and approaches. Just as an example, recent work by Bazerman (1995) and by Swales (1998) is highly discourse analytic in its approach and thus more broadly interested in global features of the specialized discourse.

Basic Distinctions: Relations between Communicators

In their essence, all approaches to the study of domain related discourse are sociological in their basic assumptions. This tenet is reflected in the fact that degree of specialization of a text and, connected to that, the relation between senders and receivers in a communicative situation concerning their respective levels of expertise is traditionally seen as a crucial factor when setting up models for describing specialized communication. The basic assumption lying behind such models is that the characteristics of the participants in the communication and the purposes pursued by them are main determiners of the way texts are written. Many different stratified classifications of the relevant communicative settings have been proposed. The following tripartite stratification gives a fairly good overall picture of the factors underlying the different proposals:

- Scientific discourse: This kind of specialized discourse is characterized by the fact that both senders and receivers are experts (expert–to–expert–communication). Texts show a high degree of abstraction and a considerable amount of often standardized terminology. These characteristics are seen as consequences of the purpose of communication in such settings, viz., to develop and refine the general knowledge of a domain. Furthermore, the high prestige connected to such discourse and its highly public character means that, apart from aspects of necessary precision, stylistic features play a very important role.
- Practically oriented discourse: Here, too, we find communication where both senders and receivers are experts (expert–to–expert–communication). But in this type of discourse, the experts are not working on developing or refining the general scientific knowledge of a domain. Instead, they are solving practical problems in their daily work and use communicative devices relevant for these purposes. As an example, the difference between the two levels in the stratification is the difference between chemical scientists writing learned articles for scientific journals and chemical scientists working in the laboratory, running their experiments and collecting the data. In the first setting, the purpose is to textually present and develop new knowledge in the field in a public environment (scientific discourse); in the second setting, the purpose of the communication is to solve practical problems occurring in the daily work (practically oriented discourse). The second setting tends to be more informal and to show only the degree of, e.g., terminological specialization in its oral or written texts that is necessary for coping. Aspects of necessary precision play a major role here, although also in this stratum the sociological aspect of showing by way of the applied language that the communicative party belongs to the relevant peer group should not be neglected.
- Discourse of popular science and domain-oriented didactics: The sender-receiver relations are here always asymmetric (i.e., sender has a higher degree of relevant knowledge than the receiver(s)).

Two prototypical cases are communication between experts and consumers (e.g., in manuals) and between experts and novices (e.g., in textbooks). The purpose of this kind of communication is to convey structured knowledge of the domain to receivers who lack this knowledge, but need it for (often) practical purposes. A third prototypical case is the communication of domain-specific knowledge in magazines for popular science. This kind of communication is slightly different from the two cases first mentioned, as it is characterized by a very important element of entertainment, whereas, e.g., manuals and textbooks are more directed toward actually enhancing the receivers' knowledge in practically relevant areas. Although texts of all three prototypical kinds show lots of similarities due to the fact that sender-receiver relations are always asymmetric, they also show important differences, mainly due to the last-mentioned difference.

Traditional Approaches

In order to describe the very complex object of study, that is text from specialized domains in all their aspects, LSP linguistics has included a wide variety of disciplines and approaches from the general toolbox of linguists, see Figure 1.

In the following list, we outline some of the central approaches of LSP linguistics and relate them to the above model.

- Terminology and lexicography: A very central area of LSP linguistics and historically the first area of studies of special domain language to be developed in large scale is terminology. Traditionally, terminology has concentrated upon standardizing the use of words (lexis) and the denomination of concepts, in order to shape languages for specific purposes, so that they are optimal for scientific or professional communication. Recently, terminology has expanded toward other areas like organization, representation and handling of information and knowledge, thus contributing to the development of systems and databases for storing knowledge. Consequently, the interface between knowledge stored in words and knowledge stored in other semiotic systems is of major importance for modern terminological research. Especially technical domains have been the object of terminological studies. Another area connected to lexis and to the manageable representation of words and meanings from specialized discourse is lexicography. The main purpose of this branch of LSP linguistics is to create a sound basis for the production of specialist dictionaries that are highly functional for dictionary users. Consequently, research in this area concentrates on the user and his/her needs and the functions of dictionaries that may be deducted from such studies. Studies from this perspective have been performed in many domains, among them especially the areas of science, law, and administration, where there is a substantial need for translational tools. Finally, it is relevant to mention that research in these fields have treated in some detail the question of the importance of cultural context for the term systems of a domain. A general distinction is made between culture-dependent and culture-independent domains. A prototypical example of a culture-dependent domain is the domain of law. Term systems in

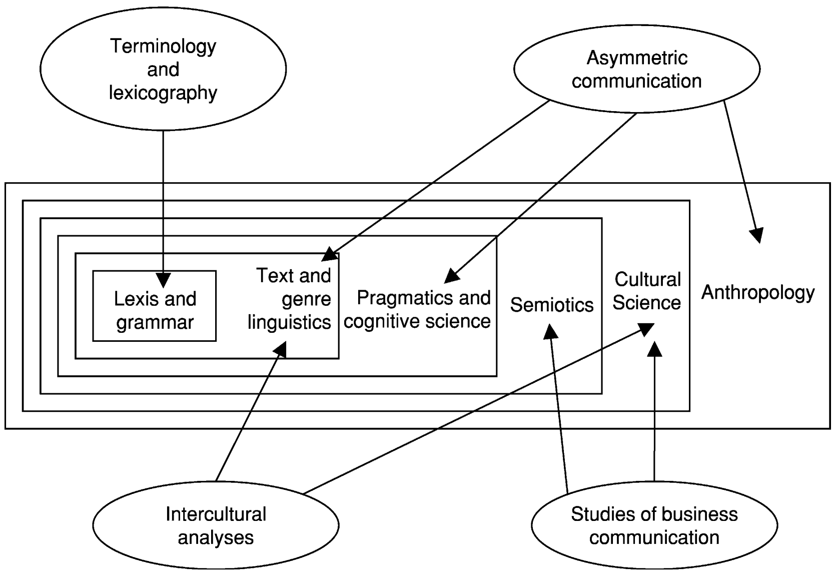


Figure 1 Linguistic disciplines with special importance for LSP linguistics.

this kind of domain are basically different across national cultures, as each nation has developed their own legal system. Consequently, legal terms from different cultures in principle do not mean the same (at least as long as no standardization has taken place). A prototypical example of a culture-independent domain is the field of electricity. In this domain, the degree of overlap between term systems from different national cultures is very much higher than in the culture-dependent domains.

- Intercultural analyses (genre analysis, register analysis, conversational analysis): The wish to develop LSP linguistics toward a more global description of discourse in specialized settings led directly to a focusing on the text as the central object of study. And in this connection, the concepts of genre and of registers have acquired paramount importance. This emphasis is due to the fact that both concepts as descriptive tools are highly oriented toward describing the influence of communicative purposes (genre) and situational settings (genre, register) on texts. Thus, to a large extent, the development of genre and register linguistics has been driven by the interest in developing descriptive tools for LSP research. Concrete studies have concentrated upon conventional characteristics of domain specific genres like research articles, legal judgments, statutes, company brochures, or scientific journals. Especially in a continental European context such studies have been performed as contrastive text and genre analyses across different cultures, thus generating results with special relevance for translators of these specialized genres. Domains with special interest have been the areas of law, medicine, and areas of (technical) science. An area where much work was produced especially in the 1990s was the area of scientific communication, and the applied perspective was the intercultural. This stress was, to a large degree, due to the fact that in these years the development toward English as lingua franca in scientific contexts was speeded up decisively. This development made it especially interesting to investigate whether in scientific contexts the national culture or the (international) discipline is the most influential concerning textual and stylistic choices in text production. An intermediary position was defended by scholars, claiming that academic writing is not influenced by national styles as such, but that it is possible to isolate four different styles, connected to different scientific cultures. No decisive evidence has been found for any of the positions, but the discussion showed different influences at different textual levels. And recently the instruments developed for this type of intercultural pragmatic studies of texts are used also in projects investigating the influence of English textual conventions on texts from different domains (like business communication) originally written in other languages. The intercultural perspective is also relevant for oral communication and has especially been investigated using conversational analysis in a number of cases. The question of differences in ways of negotiating, of showing disagreement, or in solving problems in professional settings has attracted special interest. Finally, it is important to state, that the criterion 'culture-dependent vs. culture-independent domain' is hardly relevant in this perspective, as opposed to what we saw when looking at the perspectives of terminology and lexicography: Also genres from culture-independent domains may differ considerably across national cultures.
- Asymmetric communication: As mentioned above, asymmetric discourse (i.e., communicative situations in which sender and receiver have different levels of knowledge concerning the domain of the communication) is an important part of the object of study of LSP linguistics. And this part of the object has been intensely investigated over the last 20–30 years, applying especially research instruments from text and genre linguistics, pragmatics and cognitive science, but also some instruments from anthropology, like conversational analysis. Studies are geographically concentrated around Europe and North America. The reason could be that the problem of asymmetry in knowledge relations and consequently in (abusable) power relations has been a central issue in the general discussion in these Western societies. Frequent objects of study have been doctor-patient communication, and communication in court and in other administrative settings, but also the area of technical writing (writing efficient technical documents). Research in this area has had a strong element of prescription, of not just registering characteristics of the object of study, but explicitly of contributing to better and more equal communication also in asymmetric settings.
- Studies of Business Communication: The way we have described the object of study of LSP linguistics until now, focus has been on communicative settings that are characterized mainly by their topic, by the object of communication. However, a specific branch of LSP linguistics defines its object of study more on the basis of situational elements and on the basis of the actions performed in specific texts, viz., the study of Business Communication. What makes this discourse domain specific is the fact that it is always tightly connected to the professional purposes of companies or organizations, including such areas as marketing and public relations. Typical objects of study are the internal and the external business communication within and between companies and the

management of these communicative processes. In this connection, instruments from cultural science are often applied in order to describe differences and similarities between senders and receivers from different national or international cultures. And as multimodality plays a major role in texts from these areas of communication, also the instruments from semiotics have been included and used for analyses within this sub discipline of LSP linguistics.

Recent Trends for the Development of the Discipline

In these years the recent trend in the discipline is that three areas are gradually coming into focus, thus supplementing or strengthening (some of) the perspectives and approaches already described above: cognition, semiotics, and the area of document design. The latter areas are already represented in the discipline as it looks today (cf. above), but it is likely that they will receive more attention in the future. Multi-modality is becoming a characteristic feature not only of texts from business communication, but also of texts from many other fields of specialized communication, as the importance of different media and the requirements concerning layout and entertainment are growing. So semiotics will become a necessary perspective when investigating such texts, too. And document design (connected, among other things, to the area of technical writing) has also a good future in a world, in which the amount of information is immense and the time to be allocated for reading limited. The prescriptive tradition of technical writing, intending to judge the efficiency of texts and not just collect data about the text and its coming to life, will gradually gain more importance also in fields where it has not yet been introduced, like in the areas of legal and administrative communication. And finally, there is a growing interest within the field of terminology as well as within other subfields of the discipline to focus not only upon textual or linguistic issues, but to pay more attention to the cognitive activities underlying communication in specialized settings. The expansion requires the inclusion of more instruments from the area of empirical cognitive linguistics, but is just a consequential next step in a development already running. An interesting consequence of paying more attention to this subject is that it may shift focus from superindividual aspects of communication like conventions, group language, etc. (i.e., more sociologically oriented aspects) to individual aspects like personal knowledge resources, mental processes, and learning styles. Thus, where the two prospective trends mentioned first constitute merely a profiling of the discipline within the directions in which it has been developing for the last 20 years, this last mentioned shift of attention may lead to a major paradigm shift in the discipline.

See also: Cognitive Linguistics; Communities of Practice; Conversation Analysis; Conversational Analytic Approaches to Culture; Critical Discourse Analysis; Dictionaries; Education in a Former Colonial Language; Education in a Multilingual Society; Foreign Language Teaching Policy; Formulaic Language; Functionalist Theories of Language; Genre and Genre Analysis; Human Language Processing: Connectionist Models; Human Language Processing: Symbolic Models; Institutional Talk; Language of Legal Texts; Language of Tertiary Education; Legal Genres; Legal Translation; Legal-Professional Language in Jury Trial; Lexicography: Overview; Medical Discourse and Academic Genres; Medical Discourse: Doctor–Patient Communication; Modularity of Mind and Language; Shared Knowledge; Sociolect/Social Class; Standardization; Stylistics: Pragmatic Approaches; Technical Translation; Terminology, Term Banks and Termbases for Translation; Text and Text Analysis; Visual Semiotics.

Bibliography

Bazerman C (1995). 'Systems of genre and the enactment of social intentions.' In Fredman A & Medway P (eds.) *Gene and the new thetoric.* London: Taylor & Francis. 79–101.

Bergenholtz H & Tarp S (1995). *Manual of specialised lexicography: the preparation of specialised dictionaries.* Amsterdam: Benjamins.

Bergenholtz H & Nielsen S (2002). 'Terms in the language of culture-dependent LSP dictionaries.' *Lexicographica: International Annual for Lexicography 18,* 5–18.

Bhatia V (2004). *Worlds of written discourse: a genre-based view.* New York: Continuum.

Bueno L & María R (2003). *Lenguas para Fines Específicos en España a través de sus publicaciones (1985–2002).* Madrid: Proyectos Córydon.

Bungarten T & Engberg J (2003). *Hamburger Arbeiten zur Fachsprachenforschung 5: Recht und Sprache: Eine internationale Bibliographie in juristischer und linguistischer Fachsystematik.* Tostedt: Attikon.

Clyne M (1987). 'Cultural differences in the organization of academic texts.' *Journal of Pragmatics 11,* 211–241.

Conley J M & O'Barr W M (1990). *Rules versus relationships: the ethnography of legal discourse.* Chicago: University of Chicago Press.

Fluck H-R (1996). *Fachsprachen* (5th edn.). Tübingen: Francke.

Gnutzmann C & Oldenburg H (1991). 'Contrastive text linguistics in LSP research: theoretical considerations and some preliminary findings.' In Schröder H (ed.) *Subject-oriented texts: languages for special purposes and text theory.* Tübingen: Narr. 103–136.

Göpferich S (2000). 'Analysing LSP genres (text types): from perpetuation to optimization in LSP text(-type) linguistics.' In Trosborg A (ed.) *Analysing professional genres*. Amsterdam, Philadelphia: Benjamins. 227–247.

Göpferich S (2002). *Textproduktion im Zeitalter der Globalisierung: Entwicklung einer Didaktik des Wissenstransfers*. Tübingen: Stauffenburg.

Hoffmann L (1985). *Kommunikationsmittel Fachsprache* (2nd edn.). Tübingen: Narr.

Hoffmann L, Kalverkämper H & Wiegand H-E (1998–1999). *Languages for special purposes: an international handbook of special-language and terminology research*. Berlin: de Gruyter.

Kalverkämper H (1996). 'Im Zentrum der Interessen: Fachkommunikation als Leitgröße.' *Hermes 16,* 117–176.

Kalverkämper H (1998). 'Fach und Fachwissen.' In Hoffmann L, Kalverkämper H & Wiegand H-E (eds.) 1–24.

Kleine Bibliographie fachsprachlicher Untersuchungen. Published continuously in the journal *Fachsprache. International Journal of LSP.*

Laurén C, Myking J & Picht H (1998). *Terminologie unter der Lupe: Vom Grenzgebiet zum Wissenschaftzweig.* Wien: TermNet.

Laurén C & Nordman M (1996). *Wissenschaftliche Technolekte*. Frankfurt a. M.: Lang.

Robinson P C (1991). *ESP today: a practitioner's guide.* New York/London: Prentice Hall.

Sager J P (1990). *A practical course in terminology processing*. Amsterdam: John Benjamins.

Satzger A (2001). 'Fachsprachenforschung-Akzente und Perspektiven.' *Deutsch als Fremdsprache 38(3),* 166–176.

Swales J M (1990). *Genre analysis*. Cambridge: Cambridge University.

Swales J M (1998). *Other floors, other voices: a textography of a small university building*. Mahwah, NJ: Erlbaum.

Swales J M (2000). 'Languages for specific purposes.' *Annual Review of Applied Linguistics 20, 59–76.*

Widdowson H G (1998). 'Communication and community: the pragmatics of ESP.' *ESP: English for Specific Purposes 17,* 3–14.

Languages for Specific Purposes: Pedagogy

A Johns

This article is reproduced from the Concise Encyclopedia of Pragmatics, pp. 474–480, © 1998, Elsevier Ltd.

The principal tenet of specific purposes language teaching (LSP) is that of 'specificity': each pedagogical situation and each group of learners is considered to be new and different. Ideally, a unique curriculum and collection of classroom activities is designed for each group of students and teachers in every new context. No single approach, no new or old orthodoxy, is appropriate for all pedagogical situations. Thus, central to LSP pedagogy must be the determination of the unique characteristics of each teaching/learning situation, and the exploitation of these characteristics for development of syllabuses, of classroom activities, and of evaluation and assessment.

Good LSP curricula and teaching require skills, research techniques and sensitivities that are not necessarily characteristic of 'general' language classrooms; for in LSP, the authentic world must be brought to the students, and they must learn to interact with the language as it is spoken and written in target situations. Thus, LSP teachers and curriculum designers are much more accountable than 'general' language teachers. They must be more flexible and willing to negotiate with both experts in the target situation and with the students, who are often quite familiar with the target language situation.

Specific Features of Pedagogical Contexts

Several aspects of the 'specificity' focus required in LSP are critical to curriculum development and pedagogy. First is the role of the target language in the situation in which it will be taught, i.e., the sociolinguistic context. The broadest interpretation of this principle is that an LSP course in a foreign language situation will have to be taught differently from one in a second language situation. English for Academic Purposes (EAP) in China, for example, will have considerably different objectives and classroom activities than EAP in the United States, UK, or another English-speaking country. LSPS also differ in terms of client populations and target-language features. Immigrant workers in second language situations often study the languages of the workplace (also called 'vocational languages'). In a foreign-language context, LSPS tend to be concentrated in academic or international areas, the latter of which include professional languages for negotiation and trade and the languages of the air and sea.

Related considerations include the history and politics of a target language in a pedagogical context. In North Africa, for example, the history of the competition between French and English and the political

Göpferich S (2000). 'Analysing LSP genres (text types): from perpetuation to optimization in LSP text(-type) linguistics.' In Trosborg A (ed.) *Analysing professional genres*. Amsterdam, Philadelphia: Benjamins. 227–247.

Göpferich S (2002). *Textproduktion im Zeitalter der Globalisierung: Entwicklung einer Didaktik des Wissenstransfers*. Tübingen: Stauffenburg.

Hoffmann L (1985). *Kommunikationsmittel Fachsprache* (2nd edn.). Tübingen: Narr.

Hoffmann L, Kalverkämper H & Wiegand H-E (1998–1999). *Languages for special purposes: an international handbook of special-language and terminology research*. Berlin: de Gruyter.

Kalverkämper H (1996). 'Im Zentrum der Interessen: Fachkommunikation als Leitgröße.' *Hermes 16,* 117–176.

Kalverkämper H (1998). 'Fach und Fachwissen.' In Hoffmann L, Kalverkämper H & Wiegand H-E (eds.) 1–24.

Kleine Bibliographie fachsprachlicher Untersuchungen. Published continuously in the journal *Fachsprache. International Journal of LSP.*

Laurén C, Myking J & Picht H (1998). *Terminologie unter der Lupe: Vom Grenzgebiet zum Wissenschaftzweig.* Wien: TermNet.

Laurén C & Nordman M (1996). *Wissenschaftliche Technolekte*. Frankfurt a. M.: Lang.

Robinson P C (1991). *ESP today: a practitioner's guide.* New York/London: Prentice Hall.

Sager J P (1990). *A practical course in terminology processing*. Amsterdam: John Benjamins.

Satzger A (2001). 'Fachsprachenforschung-Akzente und Perspektiven.' *Deutsch als Fremdsprache 38(3),* 166–176.

Swales J M (1990). *Genre analysis*. Cambridge: Cambridge University.

Swales J M (1998). *Other floors, other voices: a textography of a small university building*. Mahwah, NJ: Erlbaum.

Swales J M (2000). 'Languages for specific purposes.' *Annual Review of Applied Linguistics 20, 59–76.*

Widdowson H G (1998). 'Communication and community: the pragmatics of ESP.' *ESP: English for Specific Purposes 17,* 3–14.

Languages for Specific Purposes: Pedagogy

A Johns

This article is reproduced from the Concise Encyclopedia of Pragmatics, pp. 474–480, © 1998, Elsevier Ltd.

The principal tenet of specific purposes language teaching (LSP) is that of 'specificity': each pedagogical situation and each group of learners is considered to be new and different. Ideally, a unique curriculum and collection of classroom activities is designed for each group of students and teachers in every new context. No single approach, no new or old orthodoxy, is appropriate for all pedagogical situations. Thus, central to LSP pedagogy must be the determination of the unique characteristics of each teaching/learning situation, and the exploitation of these characteristics for development of syllabuses, of classroom activities, and of evaluation and assessment.

Good LSP curricula and teaching require skills, research techniques and sensitivities that are not necessarily characteristic of 'general' language classrooms; for in LSP, the authentic world must be brought to the students, and they must learn to interact with the language as it is spoken and written in target situations. Thus, LSP teachers and curriculum designers are much more accountable than 'general' language teachers. They must be more flexible and willing to negotiate with both experts in the target situation and with the students, who are often quite familiar with the target language situation.

Specific Features of Pedagogical Contexts

Several aspects of the 'specificity' focus required in LSP are critical to curriculum development and pedagogy. First is the role of the target language in the situation in which it will be taught, i.e., the sociolinguistic context. The broadest interpretation of this principle is that an LSP course in a foreign language situation will have to be taught differently from one in a second language situation. English for Academic Purposes (EAP) in China, for example, will have considerably different objectives and classroom activities than EAP in the United States, UK, or another English-speaking country. LSPS also differ in terms of client populations and target-language features. Immigrant workers in second language situations often study the languages of the workplace (also called 'vocational languages'). In a foreign-language context, LSPS tend to be concentrated in academic or international areas, the latter of which include professional languages for negotiation and trade and the languages of the air and sea.

Related considerations include the history and politics of a target language in a pedagogical context. In North Africa, for example, the history of the competition between French and English and the political

implications for studying each language influence LSP situations in public secondary schools and universities. Since most languages of wider communication have been the tongues of colonial powers or of privileged classes, the LSP curriculum is often considered to be political and is often scrutinized for cultural and political content by government officials.

The nature of the students within a specified context is also a prime consideration. Their proficiencies and their experiences with the target language as well as their first language experiences and backgrounds are central to educational planning. For instance, professionals in business, engineering and the sciences already conversant with the languages of their professions in L1 should experience a considerably different curriculum from beginning students whose understanding of domain-specific language is limited. Students who can read and write proficiently in their first languages will be enrolled in courses quite unlike those for students who are semiilliterate.

Students and professionals enter an LSP program with theories of how language is learned, e.g., including memorization, reading aloud, or drilling. These theories must also be considered in the development of teaching plans. Related to the learners' theories of language learning are their perceptions of the student/teacher role relationship. Some learners feel that the teacher is 'the boss' and should therefore be in sole control when it comes to determining what should be taught and how it should be presented. Other learners are comfortable with more consultative classroom organization, such as that found in collaborative learning, Whole Language Teaching and many communicative teaching environments.

There is a long history in LSP of examining students' needs and expectations; these are central considerations in LSP curriculum development. They can be divided into three categories: cultural–educational, personal and individual, and academic/occupational. Some LSP practitioners argue that the personal and individual factors should be the primary focus of LSP teaching, that one must start with students' present needs and interests, taking into account their anxieties and differences in learning styles, and emphasizing the pedagogical appropriateness of activities for specified learners. Others argue that with professionals and advanced academic students, in particular, academic and occupational goals, often long-term, should be the primary considerations.

Additional concerns entail the preparation, theories and proficiencies of teachers within a given pedagogical context. Teachers who have been educated in linguistics or language teaching pedagogy are often able to cope with a syllabus that is much more demanding than teachers who have been prepared in literature or another field – or who have had little preparation. Many teachers, especially in a foreign-language context, follow a single theory of language teaching such as Grammar–Translation or 'intensive reading.' Others have heard about certain methodologies and practices such as 'communicative language teaching,' but do not understand the principles and theories on which these approaches are based. Plans for LSP teaching curricula must consider teachers' backgrounds and theories; otherwise, the teachers may attempt to defeat the purposes of the curriculum designers.

In addition, there are many LSP teachers in the world whose proficiency in the target language is quite limited. Again, the choice of syllabus and language activities will have to be adapted for the language skills of the teachers, as well as for the students in the program.

Like students, teachers have visions of student/teacher role relationships. They, too, can believe that teachers are obligated to take the sole responsibility for classroom decision-making and learning. For these reasons, teachers must be an integral part of LSP curriculum development and training. If they do not accept the aims of a more learner-centered LSP program, for example, there may be conflicts between curriculum design intentions and classroom outcomes.

Another issue regarding teacher preparation and attitudes is unique to LSP: because all LSP teaching targets specific purposes settings and languages, many teachers are uncomfortable with their teaching assignments. In English for Specific Purposes (ESP), for example, the majority of the curricula deal with science and technology; however, most ESP (and LSP) teachers have been educated in the humanities. Some experts believe that the teachers' fears can be overcome when they realize that their job is to teach about language and not about content. However, others are concerned that teachers who do not understand the concepts, methods and aims of the specific purpose community (for example, of science) may become estranged and return to what they know, for example, the teaching of 'general' English from an easy-to-follow textbook.

Besides taking the various actors (i.e., teachers, students, politicians) within the pedagogical context into account, the LSP curriculum designers and instructors must study the nature of the oral and written language for which students are being prepared, and the role of the discourse identified in the sociocultural contexts in which students will be working or studying. Here, authenticity is the key to appropriate LSP pedagogy. There are differences of opinion as to the meaning of authenticity; however, in much of the world, the term refers to real, unmodified oral and

written discourse taken from the context in which students will be using the target language. In order to understand the authentic nature of discourses, or genres, pedagogies must link the language produced with the communities in which it functions. The real-world role of a text or discourse and its purpose and audience must be integral to an understanding of LSP curricula.

Because of the specificity of LSP teaching, new curricula and approaches must be constantly developed. Thus, good LSP pedagogy varies widely from country to country, from professional and occupational to academic life, and, within a specific area, from school to school and class to class. However, no approach can be isolated from other language teaching or from second language acquisition research. It must be acknowledged that there are trends in LSP teaching, influenced principally by 'general' teaching methodologies, by technology, and by language processing and learning theories. It will be the purpose of the remainder of this discussion to identify and discuss some of these trends.

Current Trends

Learner-Centered Approaches

As was mentioned previously, attention to the nature of learners is central to modern LSP teaching, as it is to much of 'general' language instruction. In its most extreme form, a learner-centered approach is based upon the principle that learning is totally determined by the nature and will of the students (Hutchinson and Waters, 1987). When discussing learner-centered pedagogies, it is essential to examine the issues of learner prior knowledge, language processing, and conceptions of tasks that influence classroom practices.

Exploiting learners' prior knowledge and experience is central to an understanding of a learner-centered pedagogy. Work in schema theory, based upon the connectionist view of language acquisition, has assisted teachers in developing pedagogies that draw upon students' knowledge of the content of the lesson (content schemata) and of the structure of the authentic discourses (formal schemata) (Carrell, 1983). When preparing for a language activity, students are given questions and activities that draw upon their often extensive first language schemata and experiences with the world, thus motivating interest and promoting familiarity with the demands of the task.

Learners also come to the classroom with specific styles for approaching learning, i.e., "cognitive, affective and physiological traits that are relatively stable indicators of how learners perceive, interact with and respond to the learning environment" (Reid, 1987: 87). Some researchers focus upon perceptual styles or learning channels: the visual, auditory, kinesthetic and tactile. Others are interested in ability to deal with ambiguity, noting that learners may be convergent, i.e., seeking closure or resolution (a style encouraged by math and sciences) or divergent, those who enjoy evaluation, criticism, and problem-solving (often students from the humanities and social sciences). Learners also have preferences for classroom organization, e.g., many of the learners in LSP contexts are field-independent, i.e., they would prefer to complete a project alone rather than in a group. In addition to styles, strategies for completing tasks are considered in a learner-centered curriculum. One of the most complete and well-developed systems of strategies appears in publications by Chamot and O'Malley (1986: 17). In their scheme, strategies can be divided into these categories:

> Metacognitive strategies: including executive processes for learning, monitoring one's comprehension and production, and evaluating how one has achieved an objective.
>
> Cognitive strategies, in which the learner interacts with the material to be learned by manipulating it mentally or physically.
>
> Social–affective strategies, in which the learner either interacts with another person in order to assist learning, or uses some kind of affective control to assist learning.

Closely related to the issue of strategies is that of language processing. From the 1980s, interest in the relationship between first and second language processing has grown, both in reading and in writing. Of particular interest are the studies of process transfer by second-language writers in specific purpose situations. For example, St John (1987) discovered that Spanish scientists writing in English were able to draw extensively from their prior knowledge of content and research article form. They were therefore able to concentrate upon their principal concerns: precise language and the development of specific article sections. These and others' insights into the processing of written texts have led to curricula that prepare LSP students for reading and writing by assisting them in activating the appropriate schemata, in drafting texts and revising, and in helping them to consider audience and purpose as they comprehend and produce text.

In learner-centered approaches, course design and teaching often become negotiated, dynamic processes, since needs, expectations, and student resources vary with each group and within a single course sequence. This does not mean that students are in total

control, by any means. However, it does suggest that LSP teachers must take into account student learning styles, strategies and language processing approaches. And, since research indicates that students benefit from expanding their strategies repertoire and from developing a metacognitive awareness of their text processing, LSP teachers should assist students in becoming more flexible and more aware of their own learning styles and approaches.

Genre-Based Approaches

Whereas some curricula begin with the nature of the learners and are designed to expand upon learner strategies and metacognition, others are sensitive to the discourse community in which the learners will be functioning, e.g., an international business, an academic department, a tourist agency, or a factory. In these genre-based curricula, the discourses from the target speech community are studied, preferably by the students themselves, in terms of text roles in the target context; for the structure, styles and purposes of discourses are defined by the communities in which they operate (see Swales, 1990).

A discourse community is a group of likeminded people with common goals and shared ways of looking at the world. Members communicate with one another through a variety of genres that serve the community's purposes, e.g., faxes, e-mail, newsletters, research articles, reports, and memos. Though discourse communities may produce the majority of genres in a single language (e.g., English) in many of the sciences, members may share values and aims that are realized in a number of different languages.

In genre-based approaches, students are immediately exposed to exemplars of genres written in the target language from within identified discourse communities. Through instruction or induction, students identify the purposes, macrostructures and linguistic features of these genres and relate their findings to the purposes of the genres within the discourse communities.

How is this done in classroom practice? One example may suffice. In academic purposes settings, students (or teachers) can act as ethnographers, interviewing experts and collecting realia from the target setting. If the community is a group of international agronomists, for example, students interview practicing agronomists to identify those genres that are central to the community, e.g., academic articles, dissertations, and newsletters. They compare various exemplars for a number of elements, e.g., macrostructure, grammar, lexicon, and metadiscoursal features. The discourses they have analyzed then become the texts for the LSP class, as students attempt to produce their own versions and to compare them with the texts by experts. Thus, in a genre-based class, the sequence of activities is (a) expert interviews, (b) collection of genres, (c) genre analysis, (d) preliminary conclusions about genres' purposes and features, (e) student genre production, and (f) evaluation by experts.

Task-Based Approaches

'Task' is – and probably will remain – ill-defined in the literature, yet the concept has become increasingly important to LSP, for it meets criterial demands and can be suitable for either a learner-centered or a genre-based approach. In learner-centered curricula, tasks are seen either as learner actions or as learner representations of tasks. 'Doing' or action tasks are related in a manner much like the behavioral objectives of the 1950s: 'Learners will use Wh-questions in controlled drills,' or 'Learners will identify the main points of a text.' Though dated, these 'doing' tasks continue to provide structure and guidance for many curricula. A more current expansion on the 'doing' concept comes from the Notional–Functional language teaching advocates. For them, 'doing' can be divided into three categories: an action sequence, a notional–functional description, and a list of teaching points. An action sequence might be 'A person arrives at a party'; the description, 'Meeting strangers, introductions, first names,' and the teaching points would be "I'm Mohamed. She's Maria. I would like you to meet my wife" (Corbel, 1985: 78).

A learner-centered task definition of the early 1990s, 'task representation,' originates in the writing process literature. In this realization, the manner in which the learners mentally represent a task which they are given is central to learner success or failure. Thus, if learners misconstrue the audience or context for a writing task, they will find it difficult to produce an acceptable written product. Likewise, if the learners misread a text because their mental representation is not consistent with that of the text writer, their comprehension will not be complete.

Others speak of abilities rather than actions as central to task completion. They list as objectives phrases such as the following: 'Students will develop the ability to assimilate new concepts and information associated with a specific subject,' 'They will be able to use linguistic resources to perform *x* task,' or 'They will understand particular modes of inquiry within a community.'

As can be seen from the last of the two definitions discussed, 'task' has become an important term both to learner processing studies and to understanding the demands of the target discourse community and its

genres. Doyle (1979) has developed an appealing, global definition to guide pedagogues:

> The term 'task' focuses attention on three aspects of students' work: (a) the products (or genres) students are to formulate, (b) the operations that are to be used to generate the product, i.e., the process, and (c) and the givens, the resources available to students in accomplishing the tasks (Doyle, 1979: 163).

These are some of the definitions of task that are influencing LSP teaching and curriculum development. Because of the plethora of definitions and because there is no agreement upon task difficulty or breakdown of tasks into mini-moves, introducing and sequencing of tasks in a curriculum is problematic. Nonetheless, task-based approaches, considered more authentic and holistic than national/functional or other alternatives, have been embraced by key figures in the LSP movement.

Technological Influences

New definitions of literacy have begun to appear in the literature which will have a major impact in the twenty-first century upon the understanding of LSP teachers' roles and instructional aims. The computer is both widely available and can be used for many different purposes, including word processing. Because of this, it has become a necessary literacy tool in the lives of many LSP students, whatever their academic or professional/vocational goals. Thus, computer literacy, involving word processing and, in some cases, programming, is a fundamental aim in many LSP courses. Car mechanics must not only find the mechanical problems, but they must input the exact terms to cue an appropriate output. Businesspeople must read and exploit printouts, access bulletin boards, and make use of networks to communicate with their colleagues worldwide.

The technological demands of LSP settings include as basic a skill as typing and advanced skills in communication and programming. To meet these demands adequately, many LSP settings are computer labs in which students network to collaborate or peer review their writing, or use the computers to communicate with members of worldwide discourse communities.

Evaluation of Curriculum and Teaching

In LSP, evaluation leads to change. The inadequate LSP curriculum is one that is overwritten or inflexible; the appropriate curriculum provides for change and input by all of those involved in its success. Because the curriculum must be specific to students' needs and their target discourse communities, teachers and students continually review the goals and purposes of their courses. Students are given opportunities to reflect upon the value of a particular classroom activity or text, in terms of their own learning styles and processes and their understanding of program goals. They can be asked, for example, whether they understood the purposes of what was being taught, whether they already know how to do an activity, whether an activity provided appropriate variety or focus, whether it assisted them in gaining understanding of the tasks they must perform, and whether they believe that they could transfer the task or use the texts in real world settings.

Teachers can also keep reflective evaluations, journals of what went well and what did not. They might ask themselves some of the following questions as they are involved in curricular revision: 'Were the needs assessments accurate and complete?' 'Were the needs assessments accurate and complete?' What do we still need to find out about students or about the target community?' 'Are the materials, methods, and activities appropriate to the newest assessment of student needs?' 'What changes should be made?' 'Are resources for teachers and students appropriate?' 'Where can we get additional resources?' 'Are learner strategies adequate?' 'What alternative strategies should we be encouraging?' 'Do learners attend regularly, apply themselves, and practice outside of class?' 'If not, why not?'

Student Assessment

Student assessment is another issue that is central to the success of a curriculum, especially in ESP programs, which are highly accountable. Though testing is an area of LSP that has been generally neglected, much can be done to demonstrate student abilities and knowledge. One important area of assessment, especially in learner-centered curricula, is of self. Several types of self-assessment measures are suggested in the literature, depending upon the goals of the assessment process. There are, for example, selfassessments of writing or reading, in which students are asked to reflect upon the success of their approaches to particular texts. There are attitude and motivational assessments as well. Peer assessment is also central to the learning process; students can be asked to consider criteria for discourses, then to evaluate the written or spoken work of their colleagues.

More traditional tests are for placement, proficiency and achievement, all of which are somewhat suspect within the LSP community. There are a number of reasons for these suspicions, the principal ones

being that it is very difficult to test students on the real-world tasks that they will be performing in discourse communities for these tasks are very complex and task-graders are generally not target community experts. Nonetheless, tests do exist, principally because it is necessary to demonstrate student progress to clients and administrators. Perhaps the best-known and most universal LSP examinations are for English: the American TOEFL which is designed to predict general academic success, and the British Council ELTS modular test, in which various forms are available for students in a number of academic disciplines (e.g., physical, social, medical sciences).

Because specific purpose language teaching requires specific purpose testing, however, the best assessments take place at the local level. One promising area of assessment comes from the Portfolio Movement, an important contribution to the teaching of writing. Student portfolios, developed over time, contain student- and teacher-selected exemplars of work that is central to course objectives. Throughout the course, students reflect upon their portfolio entries, another method for self-assessment and evaluation of growth. One example of portfolio for general academic purposes students is this: through questionnaires and interviews with subject specialists, students and faculty determine the generic tasks that appear to cross disciplines and therefore to be required in some form by all the faculty, e.g., evaluative summaries, note-taking, critical responses to readings, and analysis of data. Then throughout the LSP course, students complete these tasks and select and reflect upon task examples in their portfolios. Teachers evaluate these tasks using specific, task-based criteria. When the term is over, the students have evidence for the academic faculty of their ability to perform the generic tasks assigned.

Conclusion

Hutchinson and Waters, in their influential work (1987), speak of ESP (and LSP) as chance-taking approaches, the 'Wild West of ELT.' They also claim that LSP is the most demanding of language-teaching experiences for everyone involved. Because there are no orthodoxies as guides, teachers are faced with evolving curricular challenges. Because the teacher is obligated to prepare the students for the real world, there are new realms of knowledge and skills that must be dealt with and individuals from target discourse communities (e.g., subject specialists) who must be consulted.

Though this article has touched upon some current trends in LSP, there are many LSP courses throughout the world that do not reflect these trends. In many foreign language settings, such as in China and Japan and in parts of Latin America, there continues to be a strong emphasis upon analysis of microfeatures of texts without concern for the discourse communities in which the texts are written. Pedagogies result that are more typical of the 1960s than of the late 20th century (Swales, 1985). In many parts of the world, learner-centered methodologies are not deemed appropriate, either by the political/educational leadership or by the students. Though the understanding of technology is vital in many areas of the developed world (e.g., in the European Community), its use has barely touched other regions. In some learning contexts, subject or professional specialists are easily accessible. In others, teachers with an LSP mandate are working in relative isolation.

Nonetheless, the ideal of LSP remains: a pedagogy must consider student styles, strategies, expectations and needs seriously. It must consider, as well, the target language use contexts, in order to develop a syllabus and teaching practices that are specific to a new learning situation.

See also: Languages for Specific Purposes.

Bibliography

ALSED-*LSP Newsletter.* The LSP Centre, Dalgas Have 15, DK-20000, Denmark: Copenhagen.

Belanoff P & Dickson M (1991). *Portfolio grading.* Portsmouth, NH: Heinemann/Boynton-Cook.

Carrell P L (1983). 'Some issues in studying the role of schemata, or background knowledge, in second language comprehension.' *Reading in a Foreign Language 1,* 81–92.

Chamot A U & O'Malley J M (1986). *The cognitive academic language learning approach: An ESL content-based curriculum.* Rosslyn, VA: National Clearinghouse for Bi-lingual Education.

Corbel C (1985). 'The action sequence approach to course design.' *Prospect* (Australia) *1,* 75–90.

Doyle W (1979). 'Classroom tasks and students' abilities' In *Review of research in teacher education, 5.* Itasca, IL: Peacock. 163–198.

English for Specific Purposes: An International Journal [Formerly *ESP Journal*]. Pergamon Journals, New York.

ESP in the Classroom: Practice and Evaluation 1988 ELT Documents 128. Modern English Publications in association with the British Council, Oxford.

Fachsprache [LSP Journal]. Wilhelm Braumuller, Vienna.

Hutchinson T & Waters A (1987). *English for specific purposes: a learner-centred approach.* Cambridge: Cambridge University Press.

Johns A M (1991). 'English for Specific Purposes (ESP): Its history and contributions' In Celce-Murcia M (ed.)

Teaching English as a second or foreign language, 2nd edn. New York: Newbury House.
Kaplan R B (ed.) (1986). *Annual review of applied linguistics,* 7 (Languages and the Professions). New York: Cambridge University Press.
Kennedy C & Bolitho R (1984). *English for specific purposes.* London: Macmillan.
Nunan D (1988). *The learner-centered curriculum.* Cambridge: Cambridge University Press.
Reid J M (1987). 'The learning style preferences of ESL students.' *TESOL Quarterly 21(1),* 87–111.
Richards J C (1990). *The language teaching matrix.* Cambridge: Cambridge University Press.
Robinson P C (1979). *ESP (English for Specific Purposes). The Present Position.* Oxford: Pergamon Press.
St John M J (1987). 'Writing processes of Spanish scientists publishing in English.' *English for Specific Purposes 6(2),* 113–120.
Swales J M (1955). *Episodes in ESP.* Oxford: Pergamon Press.
Swales J M (1990). *Genre analysis. English in academic and research settings.* Cambridge: Cambridge University Press.
Tickoo M L (ed.) (1988). *ESP: The state of the art.* Anthology Series **21**, Singapore: SEAMEO Regional Language Centre.
Williams R, Swales J & Kirkman J (eds.) (1984). *Common ground: shared interest in ESP and communication studies.* ESL Documents Series **117**. Oxford: Pergamon Press (and the British Council).

Languages of Wider Communication

J Brutt-Griffler, State University of New York, Buffalo, NY, USA

Languages used to facilitate cross-linguistic communication among speakers of different languages brought together by processes of migration, trade, or travel, stretch back to the beginnings of recorded history. Some had extensive reach – covering much of the world then tied together by existing trade routes, such as the Silk Road, a loose network that linked the Mediterranean and East Asia – such as Soghdian (a language still spoken in present-day Tajikistan), and, at a later date, Persian (Foltz, 2000). But no language down to the past century could claim large numbers of speakers across the globe. For instance, while Latin served as a lingua franca in Europe for centuries, it had effectively no presence throughout Asia and Africa. One of the hallmarks of the most recent stage of *globalization* has been that the development of such a world language has become for the first time a possibility, and that English appears to have already attained that status or to be well on the way to doing so. Perhaps because the phenomenon is so comparatively new, while languages serving large expanses that formerly made up large segments of the world market, such as Persian or Latin, have been around so long, terms such as *world language* and *international language* – even at times *language of wider communication* – have tended to be employed, if not exactly interchangeably, then without clear demarcation. In their precise meanings, a *world language* must by its nature encompass the entire globe, while a language can be said to be *international* if it serves clusters of nations. For its part, a *language of wider communication* signifies one that provides a mutually intelligible medium for speakers in multilingual societies.

World Language

A *world language* is not simply the most widely spoken language in the world, or the one that is the official language in the greatest number of nations. To merit such classification, a language must have achieved a position of global preeminence in another key respect: its existence must have become a practical necessity to fill a wide range of functions brought to the fore by processes of globalization. The development of world language, then, does not represent the culmination of a linguistic process at all; it comes into being as the concomitant of larger historical trajectories, powerful forces that condition communication needs around the world. For most of the history of the world, there has been only the most remote contact between peoples in distant parts of the world. While a world market loosely existed from ancient times, the limitations of transportation and communication technology determined supranational regional zones as the largest effective units for which a lingua franca such as Persian, Arabic, or Chinese constituted a practical necessity. The need for a world language is not felt so long as communication is largely limited to bilateral relations between nations or take place in regional settings in which a particular international language such as Arabic or Spanish suits the purpose. World language requires a

Teaching English as a second or foreign language, 2nd edn. New York: Newbury House.
Kaplan R B (ed.) (1986). *Annual review of applied linguistics,* 7 (Languages and the Professions). New York: Cambridge University Press.
Kennedy C & Bolitho R (1984). *English for specific purposes.* London: Macmillan.
Nunan D (1988). *The learner-centered curriculum.* Cambridge: Cambridge University Press.
Reid J M (1987). 'The learning style preferences of ESL students.' *TESOL Quarterly 21(1),* 87–111.
Richards J C (1990). *The language teaching matrix.* Cambridge: Cambridge University Press.
Robinson P C (1979). *ESP (English for Specific Purposes). The Present Position.* Oxford: Pergamon Press.
St John M J (1987). 'Writing processes of Spanish scientists publishing in English.' *English for Specific Purposes 6(2),* 113–120.
Swales J M (1955). *Episodes in ESP.* Oxford: Pergamon Press.
Swales J M (1990). *Genre analysis. English in academic and research settings.* Cambridge: Cambridge University Press.
Tickoo M L (ed.) (1988). *ESP: The state of the art.* Anthology Series **21**, Singapore: SEAMEO Regional Language Centre.
Williams R, Swales J & Kirkman J (eds.) (1984). *Common ground: shared interest in ESP and communication studies.* ESL Documents Series **117**. Oxford: Pergamon Press (and the British Council).

Languages of Wider Communication

J Brutt-Griffler, State University of New York, Buffalo, NY, USA

Languages used to facilitate cross-linguistic communication among speakers of different languages brought together by processes of migration, trade, or travel, stretch back to the beginnings of recorded history. Some had extensive reach – covering much of the world then tied together by existing trade routes, such as the Silk Road, a loose network that linked the Mediterranean and East Asia – such as Soghdian (a language still spoken in present-day Tajikistan), and, at a later date, Persian (Foltz, 2000). But no language down to the past century could claim large numbers of speakers across the globe. For instance, while Latin served as a lingua franca in Europe for centuries, it had effectively no presence throughout Asia and Africa. One of the hallmarks of the most recent stage of *globalization* has been that the development of such a world language has become for the first time a possibility, and that English appears to have already attained that status or to be well on the way to doing so. Perhaps because the phenomenon is so comparatively new, while languages serving large expanses that formerly made up large segments of the world market, such as Persian or Latin, have been around so long, terms such as *world language* and *international language* – even at times *language of wider communication* – have tended to be employed, if not exactly interchangeably, then without clear demarcation. In their precise meanings, a *world language* must by its nature encompass the entire globe, while a language can be said to be *international* if it serves clusters of nations. For its part, a *language of wider communication* signifies one that provides a mutually intelligible medium for speakers in multilingual societies.

World Language

A *world language* is not simply the most widely spoken language in the world, or the one that is the official language in the greatest number of nations. To merit such classification, a language must have achieved a position of global preeminence in another key respect: its existence must have become a practical necessity to fill a wide range of functions brought to the fore by processes of globalization. The development of world language, then, does not represent the culmination of a linguistic process at all; it comes into being as the concomitant of larger historical trajectories, powerful forces that condition communication needs around the world. For most of the history of the world, there has been only the most remote contact between peoples in distant parts of the world. While a world market loosely existed from ancient times, the limitations of transportation and communication technology determined supranational regional zones as the largest effective units for which a lingua franca such as Persian, Arabic, or Chinese constituted a practical necessity. The need for a world language is not felt so long as communication is largely limited to bilateral relations between nations or take place in regional settings in which a particular international language such as Arabic or Spanish suits the purpose. World language requires a

stage at which every point, or virtually every point, in the world is thrown into relation with every other, when global communications has become virtually instantaneous, and the trip from any place in the world to any other has been reduced to hours rather than weeks or months. It requires a structure of world trade that is truly multilateral – rather than one that transports goods and services within definite imperially defined channels only (as was the case during colonialism in Asia and Africa).

Such international relations throw together persons from many nations from diverse regions of the world into regular contact, and make ease of communication among them an economic imperative, requiring a recognized common linguistic medium, a world language (*see* **Intercultural Pragmatics and Communication**). Finding its basis primarily in the economic realm, world language also more and more penetrates technological, scientific, intellectual, political, and even the cultural realm, as the need for translation raises a barrier to the rapidity with which information is disseminated in the digital age. In such a case, the continued spread of a world language finds a strong impetus not only in political economy, as global trade represents an increasing share of the total, but through intellectual and cultural motives, too, as people around the globe are induced to learn it to further their wide-ranging objectives (*see* **Motivation and Attitude in Second Language Learning**).

On the other hand, a world language need not be widely spoken in every nation of the world, nor must most people in the world speak it, any more than earlier regional lingua francas such as Latin in Europe or Persian along the routes of the Silk Road were universally understood in the supranational political geographies they served. Yet, at a minimum, it would be reasonable to expect such a language to be widely spoken on every continent, to have an ever-growing number of speakers in every nation, and to have emerged as the single most important language of global commerce.

The world language would not necessarily have the most mother tongue speakers, a distinction that currently belongs to Chinese, with its 874 million speakers (Grimes, 1996). On the contrary, one of the salient characteristics of a world language, like that of a supranational lingua franca like Latin or Persian in an earlier epoch, is that the majority of its users would be non-mother tongue speakers, bilinguals for whom it represents a second language. For them, a world language serves global, as opposed to purely local, functions. It is an entry-point into global relations, and one that need not usurp or even threaten the entrenched uses of the other language(s) they speak.

Is English a World Language?

Most estimates place the total number of English speakers globally at between 500 million and 1 billion, the difference consisting mainly in how proficiency in the language is defined. The best indication of the status of English as a world language is provided by the number of people around the world who feel compelled to gain at least some knowledge of the language, irrespective of their level of ultimate attainment. For it is the motivation of hundreds of millions across the globe to learn English that gives the measure of the degree to which it has become the world language, there being no comparable phenomenon associated with any other language. The rise of English is also demonstrated by the increasing frequency with which it is being incorporated into school curricula (*see* **Language Policy in Multilingual Educational Contexts**), not only in "second language contexts" (such as India and South Africa) where it serves as a *language of wider communication* on the national scale (see below), but even in foreign language contexts (including China and Europe), though this distinction is somewhat losing its usefulness for the description of English. The entrenchment of English in the school curriculum in diverse societies throughout the world will tend to produce a greater proportion of English users globally over time.

While the nature of a world language seems to preclude the need for more than one, the possibility cannot be ruled out that another language of equal, or even greater importance, could emerge, either joining or replacing English in its position of preeminence. As Graddol (2004) has argued, rapid shifts in demography together with a restructuring of linguistic space by modern telecommunications may produce profound effects on language use by the middle of the 21st century, perhaps redefining how we think of languages (local, national, regional), and simultaneously transforming the notion of world language.

Development of World Language

There have been two primary paradigms employed to account for the linguistic effects of globalization that would give rise to a world language: modernization and linguistic imperialism. Modernization describes the spread of industrial production, technology, finance, and trade, the burgeoning of the middle class, and the birth of a consumer culture, among nations that were previously 'underdeveloped'. From this standpoint, the increasing use of a common language throughout the world simply represents the natural concomitant of the march of progress. In marked contrast, linguistic imperialism

finds in the spread of certain languages to the exclusion of others the exercise of Western, ultimately imperialist, hegemony, the extension into the cultural and linguistic realm of the political and economic control that the 'Center' has exercised over the colonial and neocolonial periphery for centuries. Both explanations take for granted that globalization, including the spread of English, represents an essentially Western-driven process, one of which the rest of the world catches up to or is incorporated into European/North American society, rather than a multi-polar process driven by no central hegemony. Brutt-Griffler (2002) shows, rather, the development of English into or toward the status of a world language to be considerably more complex, a process that includes numerous forces, economic, political, cultural, and intellectual, underlying globalization, combined with the active agency of its new speakers around the globe choosing to learn the language.

World Language and Language Contact

The development of a *world language* – as one spoken primarily by non-mother tongue speakers – should produce effects on the other languages of the world and on the world language itself. As a language of bilinguals, it is thrown into contact with languages throughout the world, leading to processes of language change. Certainly the spread of English has exerted an effect on other languages via language contact. This process ranges from the incorporation of English words into languages as diverse as German and Japanese, to the emergence of mixed varieties of other languages and English, for example, chiHarare, a mixture of Shona and English spoken in Zimbabwe.

At the same time, other languages have in turn exerted an influence on English. For English itself, entering a phase of the history of the language known as the World English phase, in which non-mother tongue speakers have proliferated until, as noted above, they now exceed the number of its mother tongue users, has produced highly noticeable changes. World English – the phase of the history of the language in which it functions as a world language (Brutt-Griffler, 2002) – has spawned *World Englishes*, often called non-native varieties of the language (Kachru, 1986) spoken in nations throughout Africa and Asia (see also World Englishes). It is contested whether there is any global standard of World English, and if so, what that standard might be. But it is clear that the growing number of speakers of new varieties of World Englishes increases their potential to exert a greater impact on international usage of the language.

Linguistic, Political, and Social Consequences of World Language

The implications of the development of a world language have been significant, ranging from concerns over its alleged connection to a loss of linguistic diversity, to questions of access to high-level proficiency in a language that often confers power and privilege to its speakers, to the use of such a language to transcend ethnic identity and traditional gender roles, to questions of the meaningfulness of the paradigm of nativeness within a language of primarily non-mother tongue users (*see* **Identity: Second Language; Nonnative Speaker Teachers**). Though all of these questions are important, none has received the attention that has been devoted to the fear that the global spread of English represents a causal factor in the endangerment of languages. Graddol (2004) notes, "Many believe English will become the world language to the exclusion of all others." It has been strongly suggested that in Africa, for instance, the spread of English "is leading to the top-down displacement of numerous other tongues" (Nettle and Romaine, 2000: 144), a contention that has been challenged by others (Mufwene, 2001, 2002). Indeed, if such were the case, we would expect to find gains in the number of native speakers of English in Africa proportionate to the loss of those of disappearing languages, something the statistics do not appear to support. Rather, as Brutt-Griffler (2002) has argued, the spread of English as a world language produces not monolingual English speakers but bilinguals (see also Bilingualism and Second Language Learning). Graddol (2004) remarks, "English will indeed play a crucial role in shaping the new world linguistic order, but its major impact will be in creating new generations of bilingual and multilingual speakers across the world."

As English is learned in many multilingual contexts at an earlier age and reinforced by exposure gained through various media, new and old (Berns and de Bot, 2004), the distinction between native and non-native speakers of the language may be increasingly blurred, as fluency in it may be used as an 'exit visa' from ethnic identity and traditional gender roles (Mazrui, 2004). The continued growth of English as a world language will prompt ever-greater concerns about access to the language for all members of society, and not simply for those who can afford to buy it, and about the economic and intellectual advantages knowledge of such a world language can confer.

International Language

Though the rapid growth of the world language over the last century has taken attention away from other

international languages, a large number continue to play a vital role in the modern world, as they have for at least past millennia and will continue to do for the foreseeable future. International languages are most typically associated with the facilitating of international communication – a function that in past centuries Latin, Persian, Greek, Sanskrit, Turkish, and French all played – and the linking of diasporas across wide geographical expanses. In the modern world, however, particularly with the emergence of a world language, their primary role seems to be developing into the maintenance of supranational economic, cultural, and, within certain limitations, ethnic zones (set off by the use of, for example, Arabic, Swahili, Chinese, Spanish, Portuguese, or French). In some cases, such as Arabic or German, these represent a lack of correspondence between ethnolinguistic identity and national boundaries, as Arabs are dispersed over more than twenty nations on two continents and German is given official or special status in eight European nations. In other cases, such as French, its existence as an *international language* reflects a colonial legacy, in which most of the thirty-four nations in which French has a special status are former African or Caribbean colonies of France (see Table 1) – nations in which generally most people do not speak French as a mother tongue. A variation of this imperial case is illustrated by languages such as Spanish, which though introduced into the Americas as an imperial language, is now the mother tongue of the majority of people in more than twenty nations in the Americas, which now dominate the Spanish-speaking world. And in yet other cases, such as Swahili, an international language may constitute the expression of emerging postcolonial nationalism that plays out in a supranational context – in this case East Africa. Though Swahili is not originally indigenous to the region and has not been the mother tongue of the vast majority of its peoples, it has at least begun to take on that function for increasing numbers. Another group of international languages, of which the largest is Chinese, are more nearly expressions of a diaspora.

Table 1 Number of nations in which international languages are given some official status

1. English	63
2. French	34
3. Spanish	23
4. Arabic	23*
5. German	8
6. Portuguese	7

Ammon, 1994, p. 1726; *Crystal, 1997a, p. 359.

The categories listed above need not be mutually exclusive. For instance, Portuguese is the official language in five African nations, though it is not for the most part the mother tongue of the majority of the peoples in those nations, and yet, like Spanish in the Americas, has become the mother tongue of the majority of the largest lusophone nation, Brazil. French itself, like German, is spoken in several European nations – as well as by a diaspora population in Canada. Russian has spread to neighboring countries partly as an imperial language learned as a second language by peoples incorporated into the Russian empire and partly via a large Russian-speaking diaspora in those nations. In other cases, like Kurdish, the language can be called international, as it crosses national borders, but only because the Kurds have been denied nationhood by the three more powerful states in whose territory their would-be homeland lies. This also brings out a significant distinction between *international* and *world*, because international is based on the original construction of the *national*, world is not.

Major International Languages: A Statistical Picture

The major international languages can be distinguished through an examination of some key statistics, such as those in Tables 1 to 4, which show the number of nations in which official status is accorded to particular languages (see Table 1), and which show the number of mother tongue speakers (see Table 2), total speakers (see Table 3), and second language speakers of selected languages (see Table 4). Tables 5 through 7 provide proportions of print publications in different languages in the sciences (see Table 5 and Table 6) and humanities (see Table 7) and Table 8 shows the percentage of web pages on the Internet (see Table 8). Statistics by themselves can be revealing, but they only tell part of the story. For example, it

Table 2 Number of mother tongue speakers of 10 languages, in millions

1. Chinese	874
2. Hindi	366
3. Spanish	358
4. English	341
5. Bengali	207
6. Arabic	202*
7. Portuguese	176
8. Russian	167
9. German	100
10. French	77

World Almanac, 1999, p. 700; *Grimes, 1996.

Table 3 Total number of speakers of 11 languages, in millions

1. Chinese	1052
2. English	700+*
3. Hindi	487
4. Spanish	417
5. Russian	277
6. Bengali	211
7. Arabic	202
8. Portuguese	191
9. Indonesian	170
10. French	128
11. German	128

World Almanac, 1999, p. 700; *approximate lower estimate, Crystal, 1997b, p. 61.

Table 4 Numbers of second language speakers of 10 languages, in millions

1. English	300+*
2. Mandarin Chinese	188
3. Indonesian	140
4. Hindi	120
5. Russian	110
6. Spanish	59
7. French	51
8. Tagalog	40
9. Urdu	40
10. German	28

World Almanac, 1999, p. 700; *Crystal, 1997b, p. 54.

Table 5 Languages of articles indexed by *Chemical Abstracts* by year, in percent

	1978	*1998*
English	62.3	82.5
Chinese	0.3	5.9
Japanese	4.7	4.5
Russian	19.5	3.1
German	5.0	1.6
French	2.4	0.5
Polish	1.1	0.3
Other	4.7	1.6

Jean Laponce, 2003, p. 60.

Table 6 Percentage of languages used in natural science publications, 1980 to 1996

	1980	*1984*	*1988*	*1992*	*1996*
English	74.6	77.1	80.5	87.2	90.7
Russian	10.8	9.2	6.9	3.9	2.1
Japanese	2.3	2.5	2.1	2.3	1.7
French	3.1	2.4	2.4	1.6	1.3
German	3.5	3.3	2.9	1.6	1.2

Ammon, 1998, p. 152.

is clear that English has no rival as an international language, strengthening its claim as the world language. French, sometimes said to be a challenger, or a "big language" (Graddol, 1997), clearly does not stand out as having the same importance on a global scale, as shown by its placement as a distant sixth as a language of publishing in Chemistry, the discipline for which the most accurate statistics are available (82.5% for English, 0.5% for French, see Table 5), or fourth and second. Respectively, in the natural sciences (90.7% to 1.3%, Table 6) and the humanities (82.5% to 5.9%, Table 7) as compiled by Ammon (1998). French also ranks only fifth in number of web pages (68.4% to 3.0%, Table 8). And while French is second to English in the number of nations in which it is an officially recognized language (63 to 34, Table 1), when we examine total numbers of second language users (Table 4), we find that French, with some 50 million estimated users, ranks not only behind English (300 million+), but also Chinese (188 million), Indonesian (140 million), Hindi (120 million), and Russian (110 million). While such numbers, however, give a very approximate measure of significance as a *language of wider communication* (see below), it does not necessarily reveal much about a language's international significance. For instance, Indonesian and Hindi are mainly confined to the national borders of Indonesia and India, while French second language speakers lie mainly outside of that nation, as the number of nations in which French serves as an official medium (34) demonstrates, as compared to that of Indonesia and Hindi (1 in each case). Taking the statistics as a whole, however, we can conclude that French is an internationally significant language within the former French colonial world, but, as the statistics on scientific publishing demonstrate, it is not a language commonly chosen when the goal is to reach a world audience.

In the same manner, the contention that German represents a major international language as demonstrated by its official status in 8 nations (Ammon, 1994), placing it behind only English, French, Spanish, and Arabic, and ahead of Portuguese, looks quite different upon further analysis. Because all of those nations lie in a concentrated portion of Europe, and given estimates of only some 28 million second language speakers, it would be more accurate to describe German as a regionally important language in central Europe, of the same type, but not on the same scale, as Arabic (official in 22 nations and with more than twice the number of speakers).

The acceptance of a language as official also reflects neocolonialism in politics that may obscure

Table 7 Percentage of languages used in humanities publications, 1978 to 1995

	1978	*1982*	*1986*	*1990*	*1995*
English	69.1	69.9	70.6	71.7	82.5
French	6.6	5.9	5.9	5.9	5.9
German	5.2	6.0	5.4	5.7	4.1
Spanish	3.6	3.6	4.0	3.8	2.2

Ammon, 1998, p. 167.

both actual distributions of language users and their increasing importance globally. For example, while European languages are often official in former colonial settings, other languages do not receive similar endorsement. Thus, while German, though spoken by only 1.5% of the population of Belgium, has official status there, Turkish is not recognized as such in Germany, though its speakers make up 2.4% of the population, and though it is recognized as such in Bulgaria. There are some 15 million Turkish speakers outside of Turkey, not far behind the total of 17.5 million German speakers outside the main German-speaking nations of Germany and Austria.

Functions of International Languages

If we were to conceive the functions of an international language to include linking diasporas to mother countries, most of the world's larger languages could be said to be international in an age of increasing transnational migrations. International languages as such rather serve more institutional functions. They generally confer a sense of political affinity among groups of nations, whether, as in the case of Arab nations and Spanish-speaking American nations, as a sort of proto-nationalism, or, as in the case of France and Belgium and their former colonies, something more along the lines of a sphere of influence with colonial or neocolonial overtones. They may also facilitate the creation of a common market, whether institutionalized or more informal, particularly but by no means exclusively in the area of cultural productions, such as the Spanish-language film and music industry in Latin America. International languages allow global telecommunications giants to beam information in television signals via satellite to large international audiences – as for instance the Arabic station Al-Jazeera. Finally, they facilitate educational ties between nations sharing a common language.

Impact of English as a World Language on the Functions and Usage of International Languages

While international languages retain their importance in linking language groups across national and

Table 8 Percentage of web pages by language in 2000

English	68.4
Japanese	5.9
German	5.8
Chinese	3.9
French	3.0
Spanish	2.4
Russian	1.9
Italian	1.6
Portuguese	1.4
Korean	1.3
Other	4.6

Jacques Maurais, 2003, p. 22.

continental boundaries, they seem to be losing importance where the goal is to communicate with a global audience. There is thus a pronounced trend toward English as the medium for the publication of scientific research (Tables 5, 6, and 7). Table 6 shows that of the articles indexed by *Chemical Abstracts* the percentage in English increased from 62.3% in 1978 to 82.5% in 1998, while the proportion in every other language except Chinese fell. The two languages that followed English in 1978 experienced particularly dramatic declines thereafter, Russian (19.5% to 3.1%) and German (5.0% to 1.6%), while French fell off from 2.4% to just 0.5%. Ammon's (1998) figures (Table 7) demonstrate the same trend the natural sciences as a whole. English shows similar dominance on the Internet (Table 8). In 2000, 68.4% of web pages were in English, with only Japanese (5.9%) and German (5.8%) registering more than 5%, followed by Chinese (3.9%), French (3.0%) and Spanish (2.4%) above the 2% level (Maurais, 2003).

Languages of Wider Communication

A language of wider communication, also known as a *lingua franca*, provides a mutually intelligible medium for speakers in multilingual societies, to some extent replicating on the intra-national scale the function of world (or international) language(s) on the global scale – although world and international languages are languages of wider communication in the broadest sense of the term. There are three major categories of languages of wider communication:

- international languages which may not be indigenous to the region, such as English, French, or Portuguese in Africa, or alternatively may be indigenous, as in the case of Swahili in East Africa and Hausa in West Africa; English as a language of wider communication in Europe (Seidlhofer and Jenkins, 2003) represents something of an intermediate case, because English is indigenous to a

portion of the region but not in continental Europe where it plays its specific role;

- languages of indigenous origin that have come to fill the role of national (or regional) languages, though they are not the mother tongues of the majority of the people, for example, Bahasa Indonesia (Indonesia), Tagalog (Philippines), or Hindi (parts of India);
- languages of local and recent origin that have arisen at least in part specifically to fill the function of a language of wider communication; often called *urban vernaculars*, they are mixed languages containing elements of local languages, and sometimes international languages such as English or French; Isicamtho (South Africa), chiHarare (Zimbabwe), and Town Bemba (Zambia) number among them.

When English or French functions in this capacity, it tends to be referred to as a *second language*, to signal that it provides many speakers of a given nation with a medium in which to communicate when confronted with nationals outside the speaker's language group. In such cases, the second language may or may not be given official status, but it is generally adopted for a wide range of societal functions: economic, political, and cultural. In the second case, the languages are always official, and are also designated as national, as they are held to embody the national aspirations of the people, as may also be the case with an international language such as Swahili that is indigenous to the region. In the case of *urban vernaculars*, they are seldom given institutional recognition of any kind, and in many cases have only begun to be distinguished as languages in their own right. In the highly fluid circumstances in which they arise, they may even become transformed into the mother tongues of at least a portion of their speakers and may also be making their way into the classroom (Childs, 1997).

When a language is spoken in only one or a few countries as a second language by many of its speakers, that constitutes evidence that it functions as a language of wider communication (*see* **Lingua Francas as Second Languages**). Two such languages are Swahili, which according to some estimates has as few as 5 million mother tongue speakers and 30 million second language users, and Bahasa Indonesia, with some 17 to 30 million native speakers as compared to 140 million non-mother tongue users. Ten major languages of wider communication, with estimates of their number of second language speakers, are given in Table 4.

Functions of Languages of Wider Communication

Languages of wider communication (*lwc*) fill different functions, from the overtly political and institutional characteristic of top-down statist language planning agendas to bottom-up processes that may even go unnoticed by policymakers. When designated as official or national languages, *lwcs* can be used to promote nationalism in multi-ethnic societies, and may be associated with particular political forces or ideological aims – a role that may be as divisive as it is unifying. Hindi in India, for example, is widely associated with Hindu nationalism and viewed with suspicion, if not hostility, by large sections of the population. At the other extreme, the *lwc*s that have arisen in urban African contexts have been the spontaneous products of the mixtures within them of large numbers of speakers of various – and often closely related and mutually intelligible – languages. Closely connected, as Mufwene (2004) has pointed out, to the encroachment of urbanization on "the function of most indigenous languages as markers of ethnic identity," these languages may be used to deemphasize ethnic identity or to signal urban identity. Despite the latter connotation, or perhaps indeed because of it, such mixed languages, including South Africa's Isicamtho, chiHarare in Zimbabwe, Lingala in Congo, Town Bemba in Zambia, and Wolof in Senegal, appear to be spreading apace to rural areas, too. Because they are largely confined to non-elites, such *lwc*s, though probably the fastest-growing languages in many places, have so far received little, if any, official recognition.

See also: Bilingualism; English: World Englishes; Ethnolinguistic Vitality; Identity and Language; Identity: Second Language; Intercultural Pragmatics and Communication; Language Adaptation and Modernization; Language Change and Cultural Change; Language Maintenance and Shift; Language Planning and Policy: Models; Language Policies: Policies on Language in Europe; Language Policy in Multilingual Educational Contexts; Language Spread; Lingua Francas as Second Languages; Literacy Practices in Sociocultural Perspective; Migration and Language; Minorities and Language; Motivation and Attitude in Second Language Learning; Multiculturalism and Language; Multilingualism: Pragmatic Aspects; Native Speaker; Nonnative Speaker Teachers; Sociolinguistic Crossing; Sociolinguistics and Political Economy.

Bibliography

Ammon U (1994). 'International Languages.' In Asher R E (ed.) *The Encyclopedia of Language and Linguistics,* vol. 4. Oxford: Pergamon Press.

Ammon U (1998). *Ist Deutsch noch internationale Wissenschaftssprache? Englisch auch fur die Lehre an den deutschsprachigen Hochschulen.* Berlin/New York: Walter de Gruyter.

Berns M & de Bot K (2005). 'English language proficiency at the secondary level: a comparative study of four European countries.' In Gnutzmann C & Intemann F (eds.) *The globalisation of English and the English language classroom.* Tubingen: Gunter Narr Verlag. 192–213.

Brutt-Griffler J (2002). *World English: a study of its development.* Clevedon: Multilingual Matters Press.

Childs G T (1997). 'The status of Isicamtho, an Nguni-based urban variety of Soweto.' In Spears A K & Winford D (eds.) *The structure and status of pidgins and creoles.* Philadelphia: John Benjamins Publishing Company. 341–370.

Crystal D (1997a). *The Cambridge encyclopedia of language* (2nd edn.). Cambridge: Cambridge University Press.

Crystal D (1997b). *English as a global language.* Cambridge: Cambridge University Press.

Foltz R (2000). *Religions of the Silk Road: overland trade and cultural exchange from antiquity to the fifteenth century.* New York: St. Martin's Press.

Graddol D (1997). *The future of English.* London: The British Council.

Graddol D (2004). 'The future of language.' *Science 303,* 1329–1331.

Grimes B F (1996). *Ethnologue: languages of the world* (13th edn.). Dallas, TX: Summer Institute of Linguistics.

Kachru B B (1986). *The Alchemy of English.* Oxford: Pergamon Press.

Laponce J (2003). 'Babel and the market: geostrategy for minority languages.' In Maurais J & Morris M A (eds.) 58–63.

Maurais J (2003). 'Towards a new linguistic world order.' In Maurais J & Morris M A (eds.) 13–36.

Maurais J & Morris M A (eds.) *Languages in a globalising-world.* Cambridge: Cambridge University Press.

Mazrui A M (2004). *English in Africa: after the Cold War.* Clevedon: Multilingual Matters Press.

Mufwene S S (2001). *The ecology of language evolution.* Cambridge: Cambridge University Press.

Mufwene S S (2002). 'Colonization, globalization and the plight of "weak" languages.' *Journal of linguistics 38,* 375–395.

Mufwene S S (2004). 'Language birth and death.' *Annual Review of Anthropology 33,* 201–222.

Nettle D & Romaine S (2000). *Vanishing voices: the extinction of the world's languages.* Oxford: Oxford University Press.

Seidlhofer B & Jenkins J (2003). 'English as a lingua franca and the politics of property.' In Mair C (ed.) *The politics of English as a world language: new horizons in postcolonial cultural studies.* Amsterdam: Rodopi. 139–157.

Langue d'oc *See:* Occitan.

Lao

D A Smyth, SOAS, University of London, London, UK

Number of Speakers and Genetic Relationship

Lao belongs to the southwestern subgroup of the Tai language family. It is the national language of Laos and the first language of approximately 70% of the population of 5.4 million. There is considerable regional variation in dialects but the Vientiane dialect is regarded as standard Lao and serves as a lingua franca for the country's many ethnic minorities. Lao dialects are spoken by a further 12–15 million people in northeast Thailand, and there are sizeable overseas Lao communities in both France and the United States.

Lao bears many very close similarities to Thai and, because of Thailand's economic and cultural dominance, most Lao people understand spoken Thai. Thai radio and television broadcasts, Thai pop songs, videos of Thai movies, and foreign movies dubbed in Thai are all popular in Laos; many educated Lao people also can read Thai. Thais, however, have more difficulty understanding Lao, partly through limited exposure to the language, and partly through a lack of desire to understand it, as they regard Lao as a language of low prestige.

Phonology and Grammar

Lao is a tonal language. Standard Lao has six tones: low, mid, high, rising, high falling, and low-falling. The vowel and consonant systems closely resemble those of standard Thai; notable difference are the existence of an

Berns M & de Bot K (2005). 'English language proficiency at the secondary level: a comparative study of four European countries.' In Gnutzmann C & Intemann F (eds.) *The globalisation of English and the English language classroom*. Tubingen: Gunter Narr Verlag. 192–213.
Brutt-Griffler J (2002). *World English: a study of its development*. Clevedon: Multilingual Matters Press.
Childs G T (1997). 'The status of Isicamtho, an Nguni-based urban variety of Soweto.' In Spears A K & Winford D (eds.) *The structure and status of pidgins and creoles*. Philadelphia: John Benjamins Publishing Company. 341–370.
Crystal D (1997a). *The Cambridge encyclopedia of language* (2nd edn.). Cambridge: Cambridge University Press.
Crystal D (1997b). *English as a global language*. Cambridge: Cambridge University Press.
Foltz R (2000). *Religions of the Silk Road: overland trade and cultural exchange from antiquity to the fifteenth century*. New York: St. Martin's Press.
Graddol D (1997). *The future of English*. London: The British Council.
Graddol D (2004). 'The future of language.' *Science 303*, 1329–1331.
Grimes B F (1996). *Ethnologue: languages of the world* (13th edn.). Dallas, TX: Summer Institute of Linguistics.
Kachru B B (1986). *The Alchemy of English*. Oxford: Pergamon Press.
Laponce J (2003). 'Babel and the market: geostrategy for minority languages.' In Maurais J & Morris M A (eds.) 58–63.
Maurais J (2003). 'Towards a new linguistic world order.' In Maurais J & Morris M A (eds.) 13–36.
Maurais J & Morris M A (eds.) *Languages in a globalising-world*. Cambridge: Cambridge University Press.
Mazrui A M (2004). *English in Africa: after the Cold War*. Clevedon: Multilingual Matters Press.
Mufwene S S (2001). *The ecology of language evolution*. Cambridge: Cambridge University Press.
Mufwene S S (2002). 'Colonization, globalization and the plight of "weak" languages.' *Journal of linguistics 38*, 375–395.
Mufwene S S (2004). 'Language birth and death.' *Annual Review of Anthropology 33*, 201–222.
Nettle D & Romaine S (2000). *Vanishing voices: the extinction of the world's languages*. Oxford: Oxford University Press.
Seidlhofer B & Jenkins J (2003). 'English as a lingua franca and the politics of property.' In Mair C (ed.) *The politics of English as a world language: new horizons in postcolonial cultural studies*. Amsterdam: Rodopi. 139–157.

Langue d'oc *See:* Occitan.

Lao

D A Smyth, SOAS, University of London, London, UK

Number of Speakers and Genetic Relationship

Lao belongs to the southwestern subgroup of the Tai language family. It is the national language of Laos and the first language of approximately 70% of the population of 5.4 million. There is considerable regional variation in dialects but the Vientiane dialect is regarded as standard Lao and serves as a lingua franca for the country's many ethnic minorities. Lao dialects are spoken by a further 12–15 million people in northeast Thailand, and there are sizeable overseas Lao communities in both France and the United States.

Lao bears many very close similarities to Thai and, because of Thailand's economic and cultural dominance, most Lao people understand spoken Thai. Thai radio and television broadcasts, Thai pop songs, videos of Thai movies, and foreign movies dubbed in Thai are all popular in Laos; many educated Lao people also can read Thai. Thais, however, have more difficulty understanding Lao, partly through limited exposure to the language, and partly through a lack of desire to understand it, as they regard Lao as a language of low prestige.

Phonology and Grammar

Lao is a tonal language. Standard Lao has six tones: low, mid, high, rising, high falling, and low-falling. The vowel and consonant systems closely resemble those of standard Thai; notable difference are the existence of an

initial /ɲ/ and the absence of /r/ in spoken Lao. Grammatically, too, there are close similarities to Thai: word order is *subject-verb-object*, nouns and verbs are not inflected, the pronominal system is complex and capable of conveying subtle degrees of relative status and intimacy. 'Classifiers' or 'count words' are used in noun phrases involving numbers.

Words of purely Lao origin are often monosyllabic. Sanskrit and Pali borrowings are numerous, and where they coexist with an indigenous Lao word they reflect a more formal or literary style. Other sources of loan words are Thai, Chinese, and Cambodian, although with Thai and Lao sharing many common basic words, the extent of Lao borrowing can be overestimated; many relatively recently coined Thai words have, however, been consciously absorbed into Lao. Despite the country's former colonial status, French loan words are relatively few.

Sample of Lao with Translation

khôy	dày	hùucák	káp
1st pers. pron.	to get to	to know (s.one)	with
láaw	jūu	hóoŋhian	
3rd pers. pron.	location marker	school	

'I got to know her at school'

Recent History

When the boundaries of present-day Laos were drawn up in 1893 under the terms of a Franco-Siamese treaty, the Lao-speaking population was divided in two, the majority paradoxically being in northeast Thailand. The French brought in Vietnamese to carry out much of the administration of their colony, and with French the medium for what little postprimary education existed, the Lao language suffered a loss of prestige, even among many of its own speakers. The decline of French influence and the rise of nationalism in the aftermath of World War II helped to improve the status of Lao. Although the communist government, which came to power in 1975, has Lao-ized the education system, introduced adult literacy programmes and attempted to teach Lao to the country's ethnic minorities, literacy rates remain low.

See also: Laos: Language Situation; Southeast Asia as a Linguistic Area; Tai Languages; Thailand: Language Situation.

Bibliography

Chamberlain J R (1978). 'Language standardisation in Laos.' In Perez A Q, Santiago A O & Nguyen Dang Liem (eds.) *Papers from the conference on the standardisation of Asian languages, Manila, Philippines, December 16–21, 1974.* Pacific Linguistics, Series C – no. 47. Canberra: Research School of Pacific Studies, Australian National University.

Hoshino T & Marcus R (1981). *Lao for beginners.* Rutland, Vermont. Charles E. Tuttle.

Enfield N J (1999). 'Lao as a national language.' In Evans G (ed.) *Laos: Culture and society.* Chiang Mai. Thailand: Silkworm Books. 258–290.

Morev L N, Moskalyov A A & Plam Y Y (1979). *The Lao language.* Moscow: USSR Academy of Sciences, Institute of Oriental Studies.

Smalley W A (1994). *Linguistic diversity and national unity: language ecology in Thailand.* Chicago: University of Chicago Press.

Yates W G & Souksomboun S (1970). *Lao basic course.* Washington, D.C.: Foreign Service Institute.

Laos: Language Situation

N J Enfield, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

The Lao People's Democratic Republic (Laos) is home to languages from at least four different 'genetic' groups: Tai (southwestern and northern branches), Mon-Khmer (Bahnaric, Katuic, Vietic, Khmuic, and Palaungic branches), Hmong-Mien (Hmongic and Mienic branches), and Tibeto-Burman (Lolo-Burmese branch). The precise number of languages spoken in the country is not known, due mainly to a lack of empirical data. Estimates vary between around 70 and around 120 distinct languages (Ethnologue 2004 lists 82). For a country of its size (smaller than the United Kingdom) and population (less than 5 million), Laos features a very high degree of linguistic diversity in global terms. This diversity is greatest in the country's hill and mountain areas, which account for some 70% of the land mass. Flatter stretches of land along rivers (especially along the Mekong River) are dominated by speakers of Lao.

initial /ɲ/ and the absence of /r/ in spoken Lao. Grammatically, too, there are close similarities to Thai: word order is *subject-verb-object*, nouns and verbs are not inflected, the pronominal system is complex and capable of conveying subtle degrees of relative status and intimacy. 'Classifiers' or 'count words' are used in noun phrases involving numbers.

Words of purely Lao origin are often monosyllabic. Sanskrit and Pali borrowings are numerous, and where they coexist with an indigenous Lao word they reflect a more formal or literary style. Other sources of loan words are Thai, Chinese, and Cambodian, although with Thai and Lao sharing many common basic words, the extent of Lao borrowing can be overestimated; many relatively recently coined Thai words have, however, been consciously absorbed into Lao. Despite the country's former colonial status, French loan words are relatively few.

Sample of Lao with Translation

khôy	dày	hùucák	káp
1st pers. pron.	to get to	to know (s.one)	with

láaw	jūu	hóoŋhian
3rd pers. pron.	location marker	school

'I got to know her at school'

Recent History

When the boundaries of present-day Laos were drawn up in 1893 under the terms of a Franco-Siamese treaty, the Lao-speaking population was divided in two, the majority paradoxically being in northeast Thailand. The French brought in Vietnamese to carry out much of the administration of their colony, and with French the medium for what little postprimary education existed, the Lao language suffered a loss of prestige, even among many of its own speakers. The decline of French influence and the rise of nationalism in the aftermath of World War II helped to improve the status of Lao. Although the communist government, which came to power in 1975, has Lao-ized the education system, introduced adult literacy programmes and attempted to teach Lao to the country's ethnic minorities, literacy rates remain low.

See also: Laos: Language Situation; Southeast Asia as a Linguistic Area; Tai Languages; Thailand: Language Situation.

Bibliography

Chamberlain J R (1978). 'Language standardisation in Laos.' In Perez A Q, Santiago A O & Nguyen Dang Liem (eds.) *Papers from the conference on the standardisation of Asian languages, Manila, Philippines, December 16–21, 1974.* Pacific Linguistics, Series C – no. 47. Canberra: Research School of Pacific Studies, Australian National University.

Hoshino T & Marcus R (1981). *Lao for beginners.* Rutland, Vermont. Charles E. Tuttle.

Enfield N J (1999). 'Lao as a national language.' In Evans G (ed.) *Laos: Culture and society.* Chiang Mai. Thailand: Silkworm Books. 258–290.

Morev L N, Moskalyov A A & Plam Y Y (1979). *The Lao language.* Moscow: USSR Academy of Sciences, Institute of Oriental Studies.

Smalley W A (1994). *Linguistic diversity and national unity: language ecology in Thailand.* Chicago: University of Chicago Press.

Yates W G & Souksomboun S (1970). *Lao basic course.* Washington, D.C.: Foreign Service Institute.

Laos: Language Situation

N J Enfield, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

The Lao People's Democratic Republic (Laos) is home to languages from at least four different 'genetic' groups: Tai (southwestern and northern branches), Mon-Khmer (Bahnaric, Katuic, Vietic, Khmuic, and Palaungic branches), Hmong-Mien (Hmongic and Mienic branches), and Tibeto-Burman (Lolo-Burmese branch). The precise number of languages spoken in the country is not known, due mainly to a lack of empirical data. Estimates vary between around 70 and around 120 distinct languages (Ethnologue 2004 lists 82). For a country of its size (smaller than the United Kingdom) and population (less than 5 million), Laos features a very high degree of linguistic diversity in global terms. This diversity is greatest in the country's hill and mountain areas, which account for some 70% of the land mass. Flatter stretches of land along rivers (especially along the Mekong River) are dominated by speakers of Lao.

Historical Background

Mon-Khmer languages appear to have been spoken in Laos longer than languages of other families. They show greater geographical spread and greater internal diversity. This is particularly apparent among languages of the Vietic, Katuic, and Bahnaric subgroups spoken throughout upland areas of the country's south. Many of these languages are near extinction, with very few remaining speakers (e.g., Thémarou, with only a couple of dozen speakers). Other languages, while small, are still being learned by children (e.g., Tariang (Talieng) with about 4000 speakers), although none show the vibrancy of the northern Mon-Khmer language Kmhmu (Khmu), with some 300 000 speakers throughout northern Laos. Tai languages such as Lao have their origins in southwest migrations of Tai speakers from southwestern China, beginning some 2000 years ago (Enfield, 2003: 47–50). Like their modern descendents, these incoming Tai speakers were in search of flat riverside land ideal for their trademark 'ditch-dike' system of wet-rice cultivation. The success of Tai languages and their speakers is attested to by their present dominance of the region, with Lao the official national language of Laos (Enfield, 1999), spoken as a first language by around half the population (more than two million people). (There are also approximately 20 million residents of northeast Thailand whose native dialects are closely related to Lao.) Hmong-Mien speakers are the newest arrivals in Laos, descended from migrants coming southwards from China within the last 200 years (Culas and Michaud, 2004).

Language Contact, Sociolinguistics, Multilingualism

Laos's high degree of linguistic diversity results in intensive language contact in most parts of the country. Upland peoples maintain structured social relations across ethnicities and language backgrounds, and are normally multilingual (Bradley, 2003). By contrast, lowland-dwelling people of Lao ethnicity tend to be monolingual. Some minority languages serve as contact languages within restricted upland areas. For example, Ngkriang (Ngeq) is used in interaction among several Katuic ethnic groups in isolated Kalum district, Sekong province.

Lao is the official language of administration, education, and major economic activity, and hence all minority communities have regular contact with Lao. Minority men are often more skilled than women and children in speaking Lao due to greater contact with the language, for example, during military service. Official promotion of Lao as a national language is served in part by the development of Lao language media and education, but this is slow thanks to the country's very weak infrastructure. Authorities widely encourage (and occasionally force) migration by minority peoples from upland areas to more accessible lowlands, ostensibly to ease the burden of social development in this poor nation. Internal migration is widespread, and in many cases this accelerates the process of attrition and loss of minority languages, due to division of already fragile speech communities and their resettlement together with speakers of other languages, in closer contact with the Lao-speaking world. The result is a widespread and rapid shift to Lao.

The two most prominent and vibrant minority languages of Laos are Hmong and Kmhmu, both spoken by large populations, mainly in northern Laos. Hmong speakers in particular show no signs of abandoning their language in favor of Lao. Hmong has a roman orthography that is taught unofficially and is in widespread use in Hmong society (e.g., in advertising, development, private correspondence, etc.). This contravenes the Lao government's official stipulation that no minority language is to be written in any orthography other than the Indic-based Lao script.

International languages with a significant presence in Laos include Thai and Vietnamese. Laos's longest borders are with Thailand to the west and Vietnam to the east. Lowland communities along the Mekong River are in intensive contact with Thai via electronic media, as well as through seasonal migration of Lao workers to Thailand. Thai and Lao are essentially dialects, making Thai especially accessible to Lao speakers. Vietnamese is less widely spoken as a second language (despite large urban communities of expatriate Vietnamese in the lowlands), but is used for economic purposes by many minority peoples living along the Vietnamese border. Through the colonial period of 1893 to 1954 (Evans, 2002), French was widely used in national education and administration. The use of French as an international language has since steadily diminished in favor of English.

Linguistic Research and Reference Materials

Comparatively little scientific research has been conducted on the languages of Laos. Descriptive studies of Lao include traditional European-style grammars (e.g., Hospitalier, 1937; Royal Lao Government, 1972; Reinhorn, 1980), dictionaries (Kerr, 1972; Reinhorn, 1970; Rehbein and Sayaseng, 2000), pedagogical materials (e.g., Yates and Sayasithsena, 1970;

Hoshino and Marcus, 1981), and work of a more technical linguistic nature (e.g., Roffe, 1946; Morev *et al.,* 1972; Crisfield, 1978; Wayland, 1996; Osatananda, 1997; Enfield, 2003, 2005). Available research on minority languages includes dictionaries (e.g., Lindell *et al.,* 1994; Preisig *et al.,* 1994 [Kmhmu]; Ferlus, 1999 [Nhaheun]), text collections (e.g., Costello and IRLCS, 1993 [Katu]), grammatical descriptions (e.g., Jacq, in press [Jruq]; Enfield, 2004 [Kri]), comparative word lists (e.g., Kingsada and Shintani, 1999; Wright, 2003), and historical-coparative work (e.g., Diffloth, 1982; Peiros, 1996; Sidwell, 2000; L-Thongkhum, 2002). Field research on languages of Laos is hampered not only by difficulties of transport and communications but also by severe restrictions imposed by the socialist government on research in the social sciences.

See also: Language Change and Language Contact; Lao; Mon-Khmer Languages; Multilingualism: Pragmatic Aspects; Standardization; Tai Languages.

Language Maps (Appendix 1): Maps 105, 106.

Bibliography

Bradley D (2003). 'Language and culture of minority groups.' In Goudineau Y (ed.) *Laos and ethnic minority cultures: promoting heritage.* Paris: UNESCO. 45–68.

Costello N & Institute of Research on Lao Culture Society (1993). *Katu folktales and society.* Vientiane: Ministry of Information and Culture.

Crisfield A G (1978). *Sound symbolism and the expressive words of Lao.* Ph.D. diss., University of Hawaii.

Culas C & Michaud J (2004). 'A contribution to the study of Hmong (Miao) migrations and history.' In Tapp N, Michaud J, Culas C & Lee G Y (eds.) *Hmong/Miao in Asia.* Chiang Mai: Silkworm Books. 61–96.

Diffloth G (1982). 'Registres, dévoisement, timbres vocaliques: leur histoire en Katouique.' *Mon-Khmer Studies 11,* 47–82.

Enfield N J (1999). 'Lao as a national language.' In Evans G (ed.) *Laos: culture and society.* Chiang Mai, Silkworm Books. 258–290.

Enfield N J (2003). *Linguistic epidemiology: semantics and grammar of language contact in mainland Southeast Asia.* London: Routledge Curzon.

Enfield N J (2004). *Sketch of Kri, a Vietic language of Khammouan.* Typescript, MPI for Psycholinguistics, Nijmegen, The Netherlands.

Enfield N J (2005). 'Verbs and multi-verb sequences in Lao.' In Diller A V N & Edmondson J (eds.) *Survey of the Tai-Kadai Language family.* London: Routledge Curzon.

Evans G (2002). *A short history of Laos: the land in between.* Chiang Mai: Silkworm Books.

Ferlus M (1999). *Nhaheun-French-English lexicon.* Jacq P & Sidwell P (trans. and ed.). München: Lincom Europa.

Hoshino T & Marcus R (1981). *Lao for beginners: an introduction to the spoken and written language of Laos.* Rutland/Tokyo: Tuttle.

Hospitalier J J (1937). *Grammaire laotienne.* Paris: Imprimerie National.

Jacq P (in press). *A description of Jru'(Loven)* (2 vols). Mahidol University [Mon-Khmer Studies Special Publication].

Kerr A D (1972). *Lao–English Dictionary.* Washington, DC: Catholic University of America Press.

Kingsada T & Shintani T (eds.) (1999). *Basic vocabularies of the languages spoken in Phongsaly, Lao PDR.* Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

L-Thongkhum T (2002). 'A brief look at the thirteen Mon-Khmer languages of Xekong province, southern Laos.' In Bauer R S (ed.) *Collected papers on Southeast Asian and Pacific languages (Pacific Linguistics 530).* Canberra: Research School of Pacific and Asian Studies, ANU. 95–108.

Lindell K, Damrong T, Kingsada T & Sainyawong S (1994). *Khamu-Lao dictionary.* Vientiane: Ministry of Information and Culture.

Morev L N, Moskalev A S & Plam Y Y (1972). *The Lao language.* Moscow: Nanka (Glavnaja Redakcija Vostochnoj Literatury). [in Russian.]

Nginn P S (1984). *Elements de grammaire laotienne.* Paris: Sudestasie.

Osatananda V (1997). *Tone in Vientiane Lao.* Ph.D. diss., University of Hawaii.

Peiros I (1996). *Katuic comparative dictionary.* Canberra: Australian National University. [Pacific Linguistics Series C–132.]

Preisig E, Sayavong S & Simana' S (1994). *Kmhmu'-Lao-French-English dictionary.* Vientiane: Ministry of Information and Culture.

Rehbein B & Sayaseng S (2000). *Wörterbuch Laotisch-Deutsch/Deutsch-Laotisch.* Hamburg: Buske.

Reinhorn M (1970). *Dictionnaire Laotien-Francais.* Paris: Centre de la Recherché Scientifique.

Reinhorn M (1980). *Grammaire de la langue Lao.* Paris: Institute National des Langues et Civilisations Orientales, Université de la Sorbonne Nouvelle.

Royal Lao Government (1972). *Lao Grammar* (4 vols). Vientiane: Royal Academy, Ministry of Education [in Lao language].

Roffe E G (1946). 'The phonemic structure of Lao.' *Journal of the American Oriental Society 66,* 289–295.

Sidwell P (2000). *Proto South Bahnaric: a reconstruction of a Mon-Khmer language of Indo-China: vol. 501.* Canberra: Pacific Linguistics.

Wayland R (1996). 'Lao expressives.' *Mon-Khmer Studies 26,* 217–231.

Wright P S (2003). 'Singsali (Phunoi) speech varieties of Phongsali province.' *Language and Life Journal 1,* 62–73.

Yates W G & Sayasithsena S (1970). *Lao basic course* (2 vols). Washington, DC: Foreign Service Institute.

Laotian *See:* Lao.

Lascaris, Jean (Laskaris, Janus) (ca. 1445–1534)

N Kerecuk, London, UK

Jean Lascaris was a distinguished Byzantine scholar who, along with other Byzantine scholars, greatly influenced mainly Western European but also Eastern European thought by reintroducing original Byzantine scholarship to the mainstream of academic activity after the fall of Constantinople in 1453. He is usually described as the younger brother of Constantine Lascaris.

Jean Lascaris first came to Venice under the patronage of the Cardinal Bessarion, who sent him to Padua to study Latin. He then took refuge in Florence under the patronage of Lorenzo de Medici. Lascaris developed a series of activities to promote Hellenistic scholarship. He lectured on ancient Greeks, such as the orator Demosthenes, the historian Thucydides, the dramatist Sophocles, as well as on literary works. Twice he traveled to Greece to acquire many original Greek manuscripts. Representing Lorenzo de Medici, he negotiated with Sultan Bayezid II for the acquisition of those manuscripts. In the 1490s, the press of Lorenzo Alopa began to work with and for Lascaris in Florence. Following the expulsions of Medicis from Florence in 1495, King Charles VIII of France invited Lascaris to Paris, where he taught Greek. He held various public offices and served as the ambassador in Venice (1503–1508). Then he became a member of the New Academy of Aldus Manutius. It has been suggested that Lascaris was a friend of and influenced Manutius, but there is no acknowledgement. In 1515, Lascaris accepted the invitation of Pope Leo X to direct the Greek Academy of Rome. Francis I of France employed Lascaris with Bude to move the Blois library to Fontainbleau, creating the Royal Library, which later was to become the National Library.

Lascaris either wrote or edited a number of works and, most importantly, a number of *editiones principes,* such as *Anthologia epigrammatum graecorum cura Joannis Lascaris* (1494), *Didymi Alexandrini scholia in Iliadem* (1517), *Homericarum quaestionum liber* (1518, Porphyrius of Tyre), and *De veris Graecarum literarum formis ac causis apud antiques* (1556). The introduction of Byzantine scholarship through original Greek language became relevant to the debates of the defense of the vernacular languages first in Italy, with Alessandro Citolini (*La Lettera d'Alessandro Citolini in difesa de la lingua volgare*, 1540) and Sperone Speroni (*Dialogo delle lingue,* 1542). The latter, in an imaginary debate in which P. Bembo confronts L. Bonamico on the use of Latin, includes another dialogue opposing P. Pomponazzi and Lascaris on the value of various languages and dialects (cf., Padley, 1988: 63–64).

The works of Constantine and Jean Lascaris as well the Byzantine works circulating in Western Europe affected Eastern European thought, particularly through the Orthodox Brotherhood Schools first in Ukraine and later among the East Slavs. The Lviv Dormition Brotherhood (established in 1463) became the model for other similar schools and disseminated humanist Greek traditions through its schools and presses. The Lviv Brotherhood published *Adelphotes*, the first Greek grammar, in 1591. Greek grammatical thinking left a mark on many of the subsequent grammars, both Greek and vernacular.

See also: Language Planning and Policy: Models.

Bibliography

Davies M (2004). *Incunabula: The printing revolution in Europe, 1455–1500. Unit 45: Printing in Greek.* Primary Source Microfilm.

Geanakoplos D (1966). *Byzantine east & Latin west: Two worlds in middle ages and renaissance.* New York: Harper Torch Books.

Jordan D R (1995). *Rômaios kai filellên, romanus et graecorum studiosus: A celebration of Aldus manatius 500 years after his first dated publication.* Athens: Trochali; New Rochelle, NY: A. D. Caratzas.

Martins W (2001). *A palavra escrita. História do livro, da imprensa e da biblioteca,* 3ª edição. São Paulo: Editora Ática.

Padley G A (1988). *Grammatical theory in Western Europe 1500–1700. Trends in vernacular grammar II.* Cambridge: Cambridge University Press.

Toscano S (1998). 'Orthodox Slavedom.' In Lepschy G (ed.) *History of linguistics. Renaissance and early modern linguistics,* vol. III. London & New York: Longman. 123–148.

Lasch, Agathe (1879–1942)

J L Flood, University of London (Emeritus Professor), London, UK

Prevented by her gender and her Jewish religion from becoming an academic librarian, Agathe Lasch, born in Berlin on July 4, 1879, instead qualified as a teacher in 1898 and worked in Berlin and Halle until 1906, when, at the age of 27, she passed her *Abitur*, which would permit her to study at a university. Although women were admitted to Prussian universities at that time, in practice this depended on each professor's discretion, and the distinguished Berlin medievalist Gustav Roethe would not countenance them in his lectures or seminars. With Berlin thus closed to her Lasch studied first under the Halle phonetician and dialectologist Otto Bremer and then at Heidelberg, where in 1909 she took her doctorate under Wilhelm Braune with a thesis on the supplanting of Low German (Plattdeutsch) by High German (German, Standard) in Berlin in the 16th century (Lasch, 1910). Glowingly reviewed by Edward Schröder (*Niederdeutsches Jahrbuch 26* (1910), 151–154), her study of the linguistic interaction of the chanceries of the Brandenburg rulers with other institutions and groupings of writers would serve as an authoritative model for the investigation of linguistic development in other towns.

From 1910 to 1916 she taught German at Bryn Mawr College in Pennsylvania. Although she enjoyed her work there, once the United States was about to enter the European war, she felt that as someone who admitted to being passionate about only two things, German philology and Germany (Hartau, 2002: 11), her place was at home. Thus, in 1916 she returned to Germany to work with Conrad Borchling on a dictionary of Hamburg German. When the University of Hamburg was formally established in 1919 she was appointed *Privatdozentin*, and in 1923 she became the first woman there to be awarded the title of Professor. From 1926 to 1934 she had special responsibility for Low German. After Hitler's accession to power in 1933, the National Socialists demanded the removal of Jews from office. Despite the intervention of colleagues in Sweden and Hamburg, Lasch was forced to retire in 1934. This development, though devastating, was not unexpected. She had saved some money, and in 1937 she moved to Berlin to live with her sisters. Jews were banned from public libraries; her own library was confiscated and her pension payments ceased. Her attempts to find employment abroad came to nothing: the German government prevented her from accepting invitations she received from Dorpat in 1939 and later from Oslo. On August 12, 1942 she was arrested, deported to the '*Reichsjudenghetto*' at Riga and was never seen again. The books she was preparing, including a study of the language of Lübeck city clerks and a history of Low German, never appeared.

Lasch's lasting memorial is her work on Middle Low German. Her *Mittelniederdeutsche Grammatik* (1914), prepared at Bryn Mawr, represented a considerable broadening of her earlier studies on Berlin, covering the Low German dialects from the 13th to the 17th centuries, the heyday of the Hanseatic League. Of her *Mittelniederdeutsches Handwörterbuch* she published the first fascicles, covering A–E, between 1928 and 1934. The *Hamburgisches Wörterbuch*, begun in 1917, started appearing under the editorship of J. Meier and D. Möhn in 1985 and is expected to reach completion sometime after 2010. Lasch was particularly interested in the language of Berlin: her 1910 dissertation was followed by *Berlinisch* (1928). Her focus was on the history of the Low German language in its cultural and political context. Thanks to her, Low German studies have come to be recognized as a distinct field of enquiry, and but for her pioneering work it is unlikely that Low German would have qualified for inclusion in the European Charter for Regional and Minority Languages (1998).

Her fate resulted in an outpouring of commemorative essays soon after the war (Borchling, 1946; Korlén, 1947; Jaatinen, 1947) and again on her centenary (Peters and Sodmann, 1979; Meier, 1980). For an assessment of her work in the broader context of Jewish scholarship, see Maas (1992) and Hass-Zumkehr (2001). In 1971 a street was named after her in Hamburg ('Agathe-Lasch-Weg'), and in 1991 a triennial Agathe Lasch Prize for promising scholars was established.

See also: German; Germanic Languages.

Bibliography

Borchling C (1946). 'Agathe Lasch zum Gedächtnis: Ansprache auf der Jahresversammlung des Vereins für niederdeutsche Sprachforschung zu Goslar am 28. Sept. 1946.' *Niederdeutsche Mitteilungen 2,* 7–20.

Dalberg V & Rentenaar R (1990). 'Spuren der Bibliothek von Agathe Lasch in Dänemark.' *Niederdeutsches Jahrbuch 113,* 157–162.

Hartau C (ed.) (2002). *Zum Gedenken an Agathe Lasch (1879–1942?): Reden aus Anlass der Benennung des Hörsaals B im Hauptgebäude der Universität Hamburg in Agathe-Lasch-Hörsaal am 4. November 1999* (Hamburger Universitätsreden, neue Folge 2). Hamburg: Editiononline.de Verlag.

Hass U (2003). 'Lasch, (Luise) Agathe.' In König C (ed.) *Internationales Germanistenlexikon 1800–1950.* Berlin and New York: W. de Gruyter. II. 1160–1162.
Hass-Zumkehr U (2001). 'Agathe Lasch (1879–1942?).' In Barner W & König C (eds.) *Jüdische Intellektuelle und die Philologien in Deutschland 1871–1933.* Göttingen: Wallstein. 203–211.
Jaatinen M (1947). 'Professor Dr. Agathe Lasch zum Gedächtnis.' *Neuphilologische Mitteilungen 48,* 130–141.
Jessen M R (1947). 'Agathe Lasch, 1879–1942.' *Bryn Mawr Alumnae Bulletin,* July 1947, 7–9.
Korlén G (1947). 'Agathe Lasch et Conrad Borchling.' *Études germaniques 2,* 92–94.
Krumm K (1998). 'Agathe Lasch im Spiegel der Forschung: Widersprüche und Desiderate.' *Korrespondenzblatt des Vereins für Niederdeutsche Sprachforschung 105(2),* 58–62.
l'Aigles C de (1959). 'Agathe Lasch: aus ihrem Leben.' *Jahrbuch des Vereins für Niederdeutsche Sprachforschung 82,* 1–5.
Lasch A (1910). *Geschichte der Schriftsprache in Berlin bis zur Mitte des 16. Jahrhunderts.* Dortmund: Ruhfus. [Reprinted 1972, Wiesbaden: Sändig.]
Lasch A (1914). *Mittelniederdeutsche Grammatik.* Halle: Niemeyer [Reissued 1974, Tübingen: Niemeyer.].
Lasch A (1925). *Aus alten niederdeutschen Stadtbüchern: ein mittelniederdeutsches Lesebuch.* Dortmund: Ruhfus. [Reissued 1987 with a bibliography by Möhn, D. & Peters, R., Neumünster: Wachholtz.]
Lasch A (1928). *Berlinisch. Eine berlinische Sprachgeschichte.* Berlin: Hobbing. [Reprinted 1967, Darmstadt: Wissenschaftliche Buchgesellschaft.]
Lasch A (1928ff.). *Mittelniederdeutsches Handwörterbuch.* Hamburg: Wachholtz. [Reissued 1956, Cordes, G. (ed.), Neumünster: Wachholtz.]
Maas U (1992). 'Die vom Faschismus verdrängten Sprachwissenschaftler – Repräsentanten einer anderen Sprachwissenschaft?' In Böhne E & Motzkau-Valeton W (eds.) *Die Künste und die Wissenschaften im Exil 1933–1945.* Gerlingen: Schneider. 445–502.
Meier J (1980). 'Agathe Lasch: Worte des Gedenkens zu ihrem 100. Geburtstag.' *Jahrbuch des Vereins für Niederdeutsche Sprachforschung 103,* 7–13.
Peters R & Sodmann T (eds.) (1979). *Agathe Lasch. Ausgewählte Schriften zur niederdeutschen Philologie.* Neumünster: Wachholtz. [Reprints 17 of her articles and reviews, with a full list of her publications on pp. 487–494.].

Lasswell, Harold Dwight (1902–1978)

M Garcarz, University of Warsaw, Warsaw, Poland

Harold Dwight Lasswell (1902–1978), American sociologist, political philosopher, social psychologist, and psychoanalyst, was born on Friday the 13th, in Donnellson, Illinois. He was the son of a humble Presbyterian minister and a schoolteacher. From the very beginning, Lasswell was intensely engaged in the social matters of his times. During his school years he was an editor of the school newspaper and a pupil actively involved in school life. When he was 16, he received a scholarship to attend the prestigious University of Chicago. He graduated from there in 1922, and, in 1926, he received his Ph.D. Lasswell then attended the London School of Economics and also studied in Geneva, Paris, and Berlin, where he first had the opportunity to study psychoanalysis for sociopolitical purposes. For the next eight years he taught at the University of Chicago as a professor of social sciences. He was also a professor at Milikan University and at Columbia University, but soon after that, in 1938, Lasswell left Chicago to teach at Yale University. In 1946 he became a professor of law and political sciences and, in 1952, became a full professor of political sciences at Yale. During WW II Lasswell was a director of War Communication Research at the Library of Congress in Washington. He also taught at the John Jay College of the City of New York and was a consultant for various departments of the federal government. During 1955–1956, he was the president of the American Association of Political Sciences. In 1963 Lasswell retired from Yale University.

On the one hand, Lasswell, as a political philosopher and theoretician, was highly interested in the mechanisms that rule the contemporary political scene. On the other hand, as a social psychologist Lasswell desired to investigate the ethics and the sociocultural factors that profoundly influence those mechanisms.

Lasswell's activity as a political theoretician brought about a formulation of an assumption that politics in general is a broader concept than just a simple field of human to human relationships. He wanted to prove that the notion of politics had a more complex nature that should be investigated from the psychoanalytic angle. He created the methodological basis for the study of contemporary politics which covers the activity of humankind on many levels. At the end of the 1920s and into the early 1930s, Lasswell became an orthodox follower of Freud's ideas concerning psychoanalytical psychology, and he applied those ideas to his political science research. Later, when the concept of psychoanalysis

Hass U (2003). 'Lasch, (Luise) Agathe.' In König C (ed.) *Internationales Germanistenlexikon 1800–1950.* Berlin and New York: W. de Gruyter. II. 1160–1162.
Hass-Zumkehr U (2001). 'Agathe Lasch (1879–1942?).' In Barner W & König C (eds.) *Jüdische Intellektuelle und die Philologien in Deutschland 1871–1933.* Göttingen: Wallstein. 203–211.
Jaatinen M (1947). 'Professor Dr. Agathe Lasch zum Gedächtnis.' *Neuphilologische Mitteilungen 48,* 130–141.
Jessen M R (1947). 'Agathe Lasch, 1879–1942.' *Bryn Mawr Alumnae Bulletin,* July 1947, 7–9.
Korlén G (1947). 'Agathe Lasch et Conrad Borchling.' *Études germaniques 2,* 92–94.
Krumm K (1998). 'Agathe Lasch im Spiegel der Forschung: Widersprüche und Desiderate.' *Korrespondenzblatt des Vereins für Niederdeutsche Sprachforschung 105(2),* 58–62.
l'Aigles C de (1959). 'Agathe Lasch: aus ihrem Leben.' *Jahrbuch des Vereins für Niederdeutsche Sprachforschung 82,* 1–5.
Lasch A (1910). *Geschichte der Schriftsprache in Berlin bis zur Mitte des 16. Jahrhunderts.* Dortmund: Ruhfus. [Reprinted 1972, Wiesbaden: Sändig.]
Lasch A (1914). *Mittelniederdeutsche Grammatik.* Halle: Niemeyer [Reissued 1974, Tübingen: Niemeyer.].
Lasch A (1925). *Aus alten niederdeutschen Stadtbüchern: ein mittelniederdeutsches Lesebuch.* Dortmund: Ruhfus. [Reissued 1987 with a bibliography by Möhn, D. & Peters, R., Neumünster: Wachholtz.]
Lasch A (1928). *Berlinisch. Eine berlinische Sprachgeschichte.* Berlin: Hobbing. [Reprinted 1967, Darmstadt: Wissenschaftliche Buchgesellschaft.]
Lasch A (1928ff.). *Mittelniederdeutsches Handwörterbuch.* Hamburg: Wachholtz. [Reissued 1956, Cordes, G. (ed.), Neumünster: Wachholtz.]
Maas U (1992). 'Die vom Faschismus verdrängten Sprachwissenschaftler – Repräsentanten einer anderen Sprachwissenschaft?' In Böhne E & Motzkau-Valeton W (eds.) *Die Künste und die Wissenschaften im Exil 1933–1945.* Gerlingen: Schneider. 445–502.
Meier J (1980). 'Agathe Lasch: Worte des Gedenkens zu ihrem 100. Geburtstag.' *Jahrbuch des Vereins für Niederdeutsche Sprachforschung 103,* 7–13.
Peters R & Sodmann T (eds.) (1979). *Agathe Lasch. Ausgewählte Schriften zur niederdeutschen Philologie.* Neumünster: Wachholtz. [Reprints 17 of her articles and reviews, with a full list of her publications on pp. 487–494.].

Lasswell, Harold Dwight (1902–1978)

M Garcarz, University of Warsaw, Warsaw, Poland

Harold Dwight Lasswell (1902–1978), American sociologist, political philosopher, social psychologist, and psychoanalyst, was born on Friday the 13th, in Donnellson, Illinois. He was the son of a humble Presbyterian minister and a schoolteacher. From the very beginning, Lasswell was intensely engaged in the social matters of his times. During his school years he was an editor of the school newspaper and a pupil actively involved in school life. When he was 16, he received a scholarship to attend the prestigious University of Chicago. He graduated from there in 1922, and, in 1926, he received his Ph.D. Lasswell then attended the London School of Economics and also studied in Geneva, Paris, and Berlin, where he first had the opportunity to study psychoanalysis for sociopolitical purposes. For the next eight years he taught at the University of Chicago as a professor of social sciences. He was also a professor at Milikan University and at Columbia University, but soon after that, in 1938, Lasswell left Chicago to teach at Yale University. In 1946 he became a professor of law and political sciences and, in 1952, became a full professor of political sciences at Yale. During WW II Lasswell was a director of War Communication Research at the Library of Congress in Washington. He also taught at the John Jay College of the City of New York and was a consultant for various departments of the federal government. During 1955–1956, he was the president of the American Association of Political Sciences. In 1963 Lasswell retired from Yale University.

On the one hand, Lasswell, as a political philosopher and theoretician, was highly interested in the mechanisms that rule the contemporary political scene. On the other hand, as a social psychologist Lasswell desired to investigate the ethics and the sociocultural factors that profoundly influence those mechanisms.

Lasswell's activity as a political theoretician brought about a formulation of an assumption that politics in general is a broader concept than just a simple field of human to human relationships. He wanted to prove that the notion of politics had a more complex nature that should be investigated from the psychoanalytic angle. He created the methodological basis for the study of contemporary politics which covers the activity of humankind on many levels. At the end of the 1920s and into the early 1930s, Lasswell became an orthodox follower of Freud's ideas concerning psychoanalytical psychology, and he applied those ideas to his political science research. Later, when the concept of psychoanalysis

was no longer fully efficient as far as its application to political research was concerned, Lasswell turned to investigating the theoretical basis of democracy and the means that should be taken to uphold it.

Lasswell's academic and scientific works can be separated into two distinct periods, early and late. In the first period, during the 1930s and 1940s, Lasswell was convinced that innate logic is fully functional only when accompanied by ethic values which an individual acquires during his or her upbringing. He also claimed that according to the Freudian concept of natural instincts (e.g., Oedipal's complex), personal motives for the sociopolitical activity of politicians are shaped by their moral values, learned in childhood. In one of his first books, *Psychopathology and politics*, he explains the complexity of the socioethic processes which shape the mental stability of a fully conscious *homo politicus*. Such an individual is a unique combination of motivations and certain interpersonal skills, which, to a very great extent, influences his or her political orientation. In his research, Lasswell pays particular attention to the concept of political propaganda and the credibility of politicians as actors on the political scene. In his famous 'Models of communication,' Lasswell explains how that message, as a basic ingredient of propaganda, goes:

> *Who (says) What (to) Whom (in) What Channel (with) What Effect*

'**Who**' is the side controlling the message transfer in the media, '**says what**' is the true subject matter of the transmitted message, '**to whom**' is the receiver of the final message, '**in what channel**' stands for the means applied to transmit the message, and '**with what effect**' is the impression that the entire message transmission makes on the text receiver. For Lasswell, propaganda is a technique of modeling people's activity and shaping the attitudes of social groups toward particular political issues. Lasswell views politics as a ruthless and hermetic system of a mutual dependency of individuals who should remain loyal to the system itself. The manipulation of symbols for political purposes is, according to Lasswell, immoral from the general assumption. In the 1930s and early 1940s, Lasswell's studies of the issue of personality and politics contributed greatly to behavioral science. However, in his works from the late 1940s, there was a gradual change in Lasswell's interests.

In the second period of his scientific and public activity (the early 1950s), Lasswell was a political scientist who laid the foundation of the American theoretical cultural neopsychoanalytical theory. In *World revolutionary elites studies in coerci* (published in 1965), Lasswell elaborates his definition of elites. He claims that there are three distinct basic needs of contemporary society, which political elites must satisfy in order to uphold their real political power: prestige, income(s), and the feeling of security for fellow citizens. According to a neo-Freudian concept of the ego-psychology (the self-sense system of a person), only political scientists possess the precise instruments to monitor the political tensions within an investigated social or ethnic group, to help the elites in sustaining the political system. The experts in political sciences should take responsibility for making the political system comprehensible and transparent to voters, and universal for the ruling elites. Lasswell's approach toward politics radically changed at the end of the 1950s and into the 1960s, when he started to promote the concept of politics directed at an individual. For Lasswell, a mentally and morally sane man is the basic element of a sane community. He claimed that contemporary societies needed, e.g., 'political psychiatrists' who would prevent war and global conflicts. According to Lasswell, political stability from a global perspective resulted from the bipolar system of balancing power between the USSR and the United States.

Lasswell was also deeply interested in the concept of democratic systems. In short, he explained that democracy is the effect of global stability. On the one hand, Lasswell was of the opinion that democracy is a political system that did not satisfy the needs of individuals, but rather the community as a whole. On the other hand, he thought that political crusades that were aimed at changing the general structure of a political system brought no results unless the individuals gained economic stability. Lasswell believed that in times when traditional ethic values were diminishing, and the threat of nuclear conflict was becoming more and more real, people gradually changed their attitude toward moral issues which were the permanent cornerstones of democracy.

In his research, Laswell applied only theoretical psychoanalysis. He never practiced as a clinical specialist; therefore, his research could be treated as an incomplete or a unilateral description of human personality functioning in real politics. It is vital to highlight that Lasswell does not criticize the idea of politics, or political ideas, but the manner in which politics is served to the average voter, citizen, and human.

See also: Media and Language: Overview.

Bibliography

Lasswell H D (1975). *Environment and society in transition world.* New York: Academy of Science.

Lasswell H D (1977). *Psychopathology and politics.* Chicago: The University of Chicago Press.

Lasswell H D & Clecerland C (eds.) (1962a). *Ethics and bigness: scientific, academic, religious, political, and military.* New York: Harper & Row.
Lasswell H D & Clecerland C (eds.) (1962b). *Ethics of power: the interplay of religion, philosophy, and politics.* New York: Harper & Row.
McDougal M, Lasswell H D & Chen L-C (eds.) (1980). *Human rights and world public order. The basic policies of an international law of human dignity.* New Haven and London: Yale University Press.
Rogow A (1969). *Politics, personality, and science in the twentieth century.* Chicago: The University of Chicago Press.

Latacz, Joachim (b. 1934)

V Van Elst, Seminarium Historiographiae Linguisticae, Leuven, Belgium

Joachim Latacz is Professor Emeritus of Greek Philology at Basel University. Having specialized in early Greek epic and lyric, Greek tragedy, Alexandrian poetry, and the history of Troy, he has especially developed into a distinguished Homeric scholar. His aim to bring Homer nearer to a broad but discerning public has led to a series of captivating lectures and much-quoted articles in this field, as well as to several creditable, widely known books.

Born in Kattowitz, Latacz studied classics, ancient history, philosophy, Slavonic (Slavonic, Old Church), and comparative Indo-European linguistics at the universities of Halle (Saale), Berlin, and Hamburg (1953–1960). While collaborating on the *Lexikon des frühgriechischen Epos* (*LfgrE*) in Hamburg under the direction of B. Snell and H. Erbse (1960–1965), he obtained his doctoral degree at the Free University Berlin in 1963 with a dissertation on the semantic field 'joy' in the Homeric language (1966). After qualifying to teach at the university in Würzburg in 1972, he was appointed full professor in Mainz in 1978. Since 1981, his scholarly activities have taken place in Basel, where he was accorded emeritus status in 2002.

Among Latacz's most famous publications are his books *Homer* (1st edn. 1985) and *Troia und Homer* (1st edn. 2001), both reedited several times and translated into other languages. While the former puts the poetical and aesthetical qualities of the Homeric epopees in the forefront, the latter highlights the historical reality behind the Trojan War epic, invigorated by recent archaeological discoveries and fresh impulses from many other disciplines related to ancient history. Latacz's expertise indeed covers a whole range of research areas, varying from oriental studies to cognitive psychology. Specifically for linguistics one can mention his constant interest in the Homeric language (cf., e.g., *LfgrE* and 1966), including its formulaic, oral, and improvisatory features. In search of the historical core of the Trojan legend, he has mastered Hittite and Mycenaean too (cf., e.g., 2002).

In April 2004 Latacz celebrated his seventieth birthday. Still, his life's work of keeping Homer alive is not finished. He is currently working on his most ambitious project ever, launched in 1995: the composition of a German commentary on the *Iliad*. Because its predecessor (the edition of K. F. Ameis and C. Hentze) dates from 1868–1913 and because Homeric scholarship has since been enriched with many new insights and perspectives (including the oral-poetry theory introduced by M. Parry in 1928), Latacz's commentary will undoubtedly be a great addition to the German tradition of 'Homer-Forschung.' The commentaries on the first two books have already been published, and those on books 3, 6, 9, 19, 22, and 24 are expected in the coming years.

See also: Greek, Ancient.

Bibliography

Ameis K F, Hentze C & Cauer P (1868–1913). *Homers Ilias*. Leipzig: Teubner.
Kürschner J (2003). *Kürschners Deutscher Gelehrten-Kalender 2003. Band II: K-Scho* (19th edn.). München: Saur.
Latacz J (1966). *Zum Wortfeld 'Freude' in der Sprache Homers*. Ph.D. diss., Heidelberg: Winter.
Latacz J (ed.) (2000–). *Homers Ilias. Gesamtkommentar.* München & Leipzig: Saur.
Latacz J (2002). 'Troia-Wilios-Wilusa. Drei Namen für ein Territorium.' In Aslan R, Blum S & Kastl G (eds.) *Mauerschau: Festschrift für Manfred Korfmann,* vol. 3. Tübingen: Greiner. 1103–1121.
Latacz J (2003a). *Homer. Der erste Dichter des Abendlands* (4th edn.). Düsseldorf & Zürich: Patmos (Artemis).
Latacz J (2003b). *Troia und Homer. Der Weg zur Lösung eines alten Rätsels* (4th edn.). München: Piper.
Parry M (1928). L'épithète traditionnelle dans Homère. Essai sur un problème de style homerique. Ph.D. diss. Paris: Belles Lettres.
Thesaurus Linguae Graecae (1955ff.). *Lexikon des frühgriechischen Epos. Begründet von Bruno Snell.* Göttingen: Vandenhoeck & Ruprecht.

Relevant Website

http://www.unibas.ch/klaphil/institut/mitarb/latacz.html

Lasswell H D & Clecerland C (eds.) (1962a). *Ethics and bigness: scientific, academic, religious, political, and military.* New York: Harper & Row.
Lasswell H D & Clecerland C (eds.) (1962b). *Ethics of power: the interplay of religion, philosophy, and politics.* New York: Harper & Row.
McDougal M, Lasswell H D & Chen L-C (eds.) (1980). *Human rights and world public order. The basic policies of an international law of human dignity.* New Haven and London: Yale University Press.
Rogow A (1969). *Politics, personality, and science in the twentieth century.* Chicago: The University of Chicago Press.

Latacz, Joachim (b. 1934)

V Van Elst, Seminarium Historiographiae Linguisticae, Leuven, Belgium

Joachim Latacz is Professor Emeritus of Greek Philology at Basel University. Having specialized in early Greek epic and lyric, Greek tragedy, Alexandrian poetry, and the history of Troy, he has especially developed into a distinguished Homeric scholar. His aim to bring Homer nearer to a broad but discerning public has led to a series of captivating lectures and much-quoted articles in this field, as well as to several creditable, widely known books.

Born in Kattowitz, Latacz studied classics, ancient history, philosophy, Slavonic (Slavonic, Old Church), and comparative Indo-European linguistics at the universities of Halle (Saale), Berlin, and Hamburg (1953–1960). While collaborating on the *Lexikon des frühgriechischen Epos* (*LfgrE*) in Hamburg under the direction of B. Snell and H. Erbse (1960–1965), he obtained his doctoral degree at the Free University Berlin in 1963 with a dissertation on the semantic field 'joy' in the Homeric language (1966). After qualifying to teach at the university in Würzburg in 1972, he was appointed full professor in Mainz in 1978. Since 1981, his scholarly activities have taken place in Basel, where he was accorded emeritus status in 2002.

Among Latacz's most famous publications are his books *Homer* (1st edn. 1985) and *Troia und Homer* (1st edn. 2001), both reedited several times and translated into other languages. While the former puts the poetical and aesthetical qualities of the Homeric epopees in the forefront, the latter highlights the historical reality behind the Trojan War epic, invigorated by recent archaeological discoveries and fresh impulses from many other disciplines related to ancient history. Latacz's expertise indeed covers a whole range of research areas, varying from oriental studies to cognitive psychology. Specifically for linguistics one can mention his constant interest in the Homeric language (cf., e.g., *LfgrE* and 1966), including its formulaic, oral, and improvisatory features. In search of the historical core of the Trojan legend, he has mastered Hittite and Mycenaean too (cf., e.g., 2002).

In April 2004 Latacz celebrated his seventieth birthday. Still, his life's work of keeping Homer alive is not finished. He is currently working on his most ambitious project ever, launched in 1995: the composition of a German commentary on the *Iliad*. Because its predecessor (the edition of K. F. Ameis and C. Hentze) dates from 1868–1913 and because Homeric scholarship has since been enriched with many new insights and perspectives (including the oral-poetry theory introduced by M. Parry in 1928), Latacz's commentary will undoubtedly be a great addition to the German tradition of 'Homer-Forschung.' The commentaries on the first two books have already been published, and those on books 3, 6, 9, 19, 22, and 24 are expected in the coming years.

See also: Greek, Ancient.

Bibliography

Ameis K F, Hentze C & Cauer P (1868–1913). *Homers Ilias*. Leipzig: Teubner.
Kürschner J (2003). *Kürschners Deutscher Gelehrten-Kalender 2003. Band II: K-Scho* (19th edn.). München: Saur.
Latacz J (1966). *Zum Wortfeld 'Freude' in der Sprache Homers*. Ph.D. diss., Heidelberg: Winter.
Latacz J (ed.) (2000–). *Homers Ilias. Gesamtkommentar.* München & Leipzig: Saur.
Latacz J (2002). 'Troia-Wilios-Wilusa. Drei Namen für ein Territorium.' In Aslan R, Blum S & Kastl G (eds.) *Mauerschau: Festschrift für Manfred Korfmann,* vol. 3. Tübingen: Greiner. 1103–1121.
Latacz J (2003a). *Homer. Der erste Dichter des Abendlands* (4th edn.). Düsseldorf & Zürich: Patmos (Artemis).
Latacz J (2003b). *Troia und Homer. Der Weg zur Lösung eines alten Rätsels* (4th edn.). München: Piper.
Parry M (1928). L'épithète traditionnelle dans Homère. Essai sur un problème de style homerique. Ph.D. diss. Paris: Belles Lettres.
Thesaurus Linguae Graecae (1955ff.). *Lexikon des frühgriechischen Epos. Begründet von Bruno Snell.* Göttingen: Vandenhoeck & Ruprecht.

Relevant Website

http://www.unibas.ch/klaphil/institut/mitarb/latacz.html

Latent Semantic Analysis

P Wiemer-Hastings, DePaul University, Chicago, IL, USA

Introduction

Latent semantic analysis (LSA) is a technique for creating vector-based representations of texts that are claimed to capture their semantic content. The primary function of LSA is to compute the similarity of text pairs by comparing their vector representations. This relatively simple similarity metric has been situated within a psychological theory of text meaning and has been shown to closely match human capabilities on a variety of tasks.

This article follows the developmental path of LSA, describing its historical context, showing how it computes and uses its vector representations, and then giving examples of the theoretical and empirical support for LSA and its current research directions.

How LSA Works

LSA (originally known as latent semantic indexing) was developed for the task of information retrieval, that is, selecting from a large database of documents a few relevant documents that match a given query. Previous approaches to this task included keyword matching, weighted keyword matching, and vector-based representations based on the occurrence of words in documents. LSA extends the vector-based approach by using singular value decomposition (SVD) to reconfigure the data. The details of this process are described following, but the intuition is that there is a set of underlying latent variables that spans the meanings that can be expressed in a particular language. These variables are assumed to be independent (and therefore orthogonal in the vector space). SVD is a matrix algebra technique that essentially reorients and ranks the dimensions in a vector space.

Because the dimensions in a vector space computed by SVD are ordered from most to least significant, if some of the less-significant dimensions are ignored, the reduced representation is guaranteed to be the best possible for that dimensionality. In LSA, the typical assumption is that only the top 300 or so dimensions (out of tens or even hundreds of thousands) are useful for capturing the meaning of texts. By basing the representations on a reduced number of dimensions, words that generally occur in similar contexts have similar vectors and will, therefore, get a high similarity rating. The discarded dimensions are assumed to be the product of noise, random associations, or some other nonessential factor.

That LSA performed information retrieval better than its rival approaches is not especially surprising. What is more surprising is how well it models human behavior on a variety of linguistic tasks. Before describing these, however, the LSA technique is described in more detail.

Although there are some variations, the most common steps are these:

- Collect a large set of (domain-relevant) text and separate it into 'documents.' For most applications, each paragraph is treated as a separate document based on the intuition that the information within a paragraph tends to be coherent and related.
- Next, create a co-occurrence matrix of documents and terms. The cell in this matrix corresponding to document x and term y contains the number of times y occurs in x. A term is defined as a word that occurs in more than one document, and no stemming or other morphological analysis is performed to try to combine different forms of the same word.

If there are m terms and n documents, this matrix can be viewed as giving a representation that has an m-dimensional vector for each document, and an n-dimensional vector for each term.

- The values in each cell may be weighted to reduce the effect of common words that occur through the corpus. A common weighting method is 'log entropy,' based on information theory, which multiplies the value by its information gain.
- SVD is invoked with a parameter k, which specifies the desired number of dimensions. (In principle, the SVD would be computed with all the dimensions to create three matrices that, when multiplied together, would produce the original data, but due to the amount of memory that this would require, this is not feasible. Instead, the currently used algorithms are optimized for dealing with sparse data spaces and compute only the most significant k dimensions of the matrices.)

The result of the preceding processing is three matrices. One has a k-dimensional vector for each document, one has a k-dimensional vector for each term in the corpus, and one has the k singular values. The first two matrices define two different vector spaces that are also different from the space defined by the original matrix. The singular values can be used to transform a vector from one space to the other. The use of these matrices depends on the application.

For information retrieval, the document vectors contain the LSA representation of each document. A query is turned into a 'pseudodoc' in the document vector space by combining the vectors for the terms in the query and dividing by the singular values. Vectors are typically compared by computing the cosine between them. (Some applications use other distance metrics.) The closest vectors from the document vector space correspond to the documents that are closest in meaning to the query (according to LSA).

In most other applications, the original documents are only used for training, that is, creating the semantic space. To compare new texts, the term vectors are combined as described earlier. Here, no manipulation with the singular values is required because the vectors are compared in the term space.

For more details about the mathematical foundations of LSA, see Golub and Van Loan (1989) and Hu *et al.* (2005). For more details about the creation of LSA spaces, see Deerwester *et al.* (1990) and Quesada (2005a).

Support for LSA

Support for LSA might be said to stem from the time of World War II when Wittgenstein wrote (what was later translated as):

> . . . for a large class of cases – though not for all – in which we employ the word "meaning" it can be defined as thus, the meaning of the word is its use in language. (Wittgenstein, 1958: 20)

A large number of psychological studies have taken Wittgenstein's words to heart and shown that LSA's behavior is closely matched with that of humans, for example:

- LSA acquires words at a similar pace to human children, sometimes exceeding the number of words to which it is exposed (see Landauer and Dumais, 1997).
- LSA's knowledge of synonyms is as good as that of second-language English speakers as evidenced by scores on the Test of English as a Foreign Language (TOEFL, see Landauer *et al.*, 1997).
- LSA can tell students what they should read next to help them learn (see Rehder *et al.*, 1998).
- LSA can even interpret metaphors such as, "My lawyer is a shark" (see Kintsch, 2001).

For textual applications, LSA has another benefit besides its high correlation with human behavior. When compared with the traditional labor-intensive approach to natural language processing – developing a grammar, a lexicon, a semantic representation, and the processing engine needed to combine them – developing an LSA-based representation is quite simple. It also has the advantage of graceful degradation. If it doesn't know a word, LSA simply ignores it and bases its representation on the other words. This has led researchers to use LSA for a variety of applications, including

- Intelligent tutoring systems that allow students to enter unconstrained natural language replies to questions (see Graesser *et al.*, 2000 and Wiemer-Hastings *et al.*, 2004)
- Grading psychology essays by comparing them to pregraded essays (see Foltz, 1996)
- Evaluating summaries of documents to help teach summarization skills (Summary Street, described at http://colit.org/)
- Helping students learn to properly integrate and cite material from multiple documents (see Britt, 2005)
- Evaluating airplane landings in a flight simulator (see Quesada, 2005b).

The only applicability constraints for LSA are that the task is text based, it can be framed in terms of computing the similarity of texts, and there is an available training corpus. The tutoring systems, for example, compare a student's answer for a question to a set of expected answers. If the student's response is close enough to a good answer, then the system gives positive feedback and moves on to the next question. If the student's answer matches an expected bad answer, then the system steers the student back on track.

Issues

The research issues facing the LSA community range from the practical to the philosophical. One basic question addresses the size and substance of the training corpus. Many effective LSA applications have been developed using relatively small corpora. In one of the successful applications mentioned earlier, LSA was trained on a corpus of only a couple hundred kilobytes with 2000 word types, 30 000 word tokens, and 325 documents. In contrast, researchers at the University of Colorado have reported that they have trained LSA on a corpus containing 750 000 word types, 550 million word tokens, and 3.6 million documents.

Unfortunately, there is little hard evidence on what the ideal size of an LSA corpus might be. The current data suggests that adding extra texts is unlikely to reduce performance, so a basic rule of thumb is, the more the better.

The obvious follow-up question is, 'What kinds of text should be included in an LSA corpus?' The

common wisdom holds that the corpus should consist of texts that are relevant to the particular target task. The domain can define a sublanguage where words are interpreted in consistent ways. Furthermore, a primary concern is to achieve sufficient coverage of the words that will be encountered in the course of running the application. A domain-specific corpus will have a higher percentage of relevant words and will thus not waste its representational power on words that will not be seen by the application.

One critical objection that is raised against the LSA approach is that not only does it ignore the syntactic structure of sentences, it also even ignores word order. In other words, LSA treats a text as a bag of words. In practice, LSA does well with longer passages of words (defined by Rehder, 1998, as more than 200 words), where syntactic details may be 'washed out,' and it also does well with single words (the TOEFL test, for example), but it does not do well on single sentences as shown by Wiemer-Hastings *et al.* (1999). A variety of approaches have attempted to deal with this, including surface parsing of sentences so that the components can be compared separately with LSA and using string edit theory to infer structural relations (see Wiemer-Hastings and Zipitria, 2001; Kanejiya *et al.*, 2003; Dennis, 2005).

Another notable gap in LSA's competence is negations, either because they are omitted from the LSA training via a stop words list or simply because their widespread use throughout a corpus renders them representationally depleted. Although no satisfactory approach yet exists for dealing with negations, a possibility would be to treat them as an essentially syntactic component that can processed as described earlier.

A more fundamental question about LSA is what its dimensions mean. Because they represent latent variables, there is no clear definition. As shown by Hu *et al.* (2003), there is a high correlation between the first dimension and the frequency of occurrence of the words in the corpus. Beyond that, there are no clear answers. There is also considerable debate as to what extent LSA captures first-order co-occurrence or higher-order co-occurrence. Recent evidence from Denhière *et al.* (2005) shows that although second-order effects do occur, large changes in the similarity measure between two words can be seen when a document is added to a training corpus in which both words occur (first-order co-occurrence).

LSA's utility and correspondence with human behavior have made it a popular technique for psycholinguistic research and text processing. The issues described here (along with others, for example, 'Is LSA at all psychologically plausible?') will keep researchers busy for years to come. A collection edited by McNamara *et al.* (2005) provides considerably more detail about the current practice and research on LSA.

See also: Computational Approaches to Language Acquisition; Computational Language Systems: Architectures; Computer-Mediated Communication: Cognitive Science Approach; Content Teaching and Learning and Language; Corpus Linguistics; Discourse Parsing, Automatic; Document Retrieval, Automatic; Human Language Processing: Connectionist Models; Human Language Processing: Symbolic Models; Human Reasoning and Language Interpretation; Information Theory; Language Processing: Statistical Methods; Lexical Acquisition; Natural Language Processing: Overview; Natural Language Understanding, Automatic; Rational Analysis and Language Processing; Semantics in Psychology; Summarization of Text, Automatic; Symbolic Computational Linguistics: Overview; Text Mining.

Bibliography

Britt A, Wiemer-Hastings P, Larson A & Perfetti C (2004). 'Using intelligent feedback to improve sourcing and integration in students' essays.' *International Journal of Artificial Intelligence in Education.* Leeds: International Artificial Intelligence in Education Society. 359–374.

Deerwester S, Dumais S T, Furnas G W, Landauer T K & Harshman R (1990). 'Indexing by latent semantic analysis.' *Journal of the American Society for Information Science 41,* 391–407.

Denhière G, Lemaire B, Bellisens C & Jhean S (2005). 'A semantic space for modeling a child semantic memory.' In McNamara *et al.* (eds.).

Dennis S (2005). 'Introducing word order.' In McNamara *et al.* (eds.).

Foltz P (1996). 'Latent semantic analysis for text-based research.' *Behavior Research Methods Instruments, and Computers 28,* 197–202.

Golub G H & VanLoan C F (1989). *Matrix computations* (3rd edn.). Baltimore: Johns Hopkins.

Graesser A, Wiemer-Hastings P, Wiemer-Hastings K, Harter D & Person N & the Tutoring Research Group (2000). 'Using latent semantic analysis to evaluate the contributions of students in AutoTutor.' *Interactive Learning Environments 8(2),* 129–147.

Hu X, Cai Z, Franceschetti D, Penumatsa P, Graesser A & Louwerse M (2003). 'LSA: the first dimension and dimensional weighting.' In Alterman R & Hirsh D (eds.) *Proceedings of the 25th Annual Conference of the Cognitive Science Society.* 587–592.

Hu X, Cai Z, Wiemer-Hastings P, Graesser A & McNamara D (2005). 'Strengths, limitations, and extensions of LSA.' In McNamara *et al.* (eds.).

Kanejiya D, Kumar A & Prasad S (2003). 'Automatic evaluation of students' answers using syntactically enhanced LSA.' In *Proceedings of the Human Language Technology Conference (HLT-NAACL 2003) Workshop on Building Educational Applications using NLP.*

Kintsch W (2001). 'Predication.' *Cognitive Science 25,* 173–202.
Landauer T & Dumais S (1997). 'A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge.' *Psychological Review 104,* 211–240.
Landauer T K, Laham D, Rehder R & Schreiner M E (1997). 'How well can passage meaning be derived without using word order? A comparison of latent semantic analysis and humans.' In *Proceedings of the 19th Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Erlbaum. 412–417.
McNamara D, Landauer T, Dennis S & Kintsch W (eds.) (2005). *LSA: a road to meaning*. Mahwah, NJ: Erlbaum. passim.
Quesada J (2005a). 'Creating your own LSA space.' In McNamara *et al.* (eds.).
Quesada J (2005b). 'Understanding representation in problem solving, judgment and decision making as the creation of multidimensional space.' In McNamara *et al.* (eds.).
Rehder B, Schreiner M, Laham D, Wolfe M, Landauer T & Kintsch W (1998). 'Using Latent Semantic Analysis to assess knowledge: Some technical considerations.' *Discourse Processes 25,* 337–354.
Wiemer-Hastings P, Allbritton D, Efron J & Arnott E (2004). 'RMT: A dialog-based research methods tutor with or without a head.' In *Proceedings of the ITS2004 Seventh International Conference*. Berlin: Springer. 614–623.
Wiemer-Hastings P, Wiemer-Hastings K & Graesser A (1999). 'How latent is latent semantic analysis?' In *Proceedings of the Sixteenth International Joint Congress on Artificial Intelligence*. San Francisco: Morgan Kaufmann. 932–937.
Wiemer-Hastings P & Zipitria I (2001). 'Rules for syntax, vectors for semantics.' In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Erlbaum. 1140–1145.
Wittgenstein L (1958). *Philosophical investigations* (2nd edn.). Anscombe G E M (trans.). Oxford: Blackwell.

Laterality Across the World's Languages

J B Hellige and M M Adamson, University of Southern California, Los Angeles, California, USA

The left cerebral hemisphere is superior to the right for many aspects of speech processing and reading. This overall left-hemisphere dominance is ubiquitous across the world's languages. However, as laterality research has expanded to include a wider variety of languages, we have learned that subtle characteristics of hemispheric asymmetry that were once treated as universal may, instead, depend on specific characteristics of a language. This seems to be particularly true for what has been termed the qualitative nature of processing printed material. With this in mind, this review highlights both the overall consistency and the variability in laterality across the world's languages. Although the review is, of necessity, illustrative rather than exhaustive, it will serve to demonstrate the importance of cross-linguistic laterality studies and to suggest factors that contribute to variability. After a brief review of laterality for identifying speech sounds, we provide a somewhat more extensive review of laterality for identifying printed words and nonwords. In each case, studies using English are reviewed first to provide a context in which to understand studies with other languages.

Processing Speech Sounds

There is a good deal of converging evidence of left-hemisphere superiority for most aspects of speech processing. Though most of the evidence comes from studies of English, laterality for speech processing has been examined across a sufficient variety of languages to demonstrate the generality of left-hemisphere superiority and to suggest a few critical factors that underlie such laterality.

English

Studies of patients with unilateral brain damage, behavioral laterality studies, functional brain imaging studies, and studies using a variety of other techniques all indicate left-hemisphere superiority for identifying acoustically presented words and nonwords. For example, in a behavioral technique known as dichotic listening, two similar but different sounds (e.g., /*ba*/ and /*da*/ or /*bad*/ and /*dad*/) are presented simultaneously, one to each ear. Because input pathways from each ear project predominantly to the contralateral cerebral cortex, better identification of the stimuli from one ear is taken to reflect superior processing in the contralateral cerebral hemisphere. For English speakers processing English speech sounds, there are remarkably consistent indications of a right-ear (that is, left-hemisphere) advantage for most people.

Kintsch W (2001). 'Predication.' *Cognitive Science 25,* 173–202.
Landauer T & Dumais S (1997). 'A solution to Plato's problem: the latent semantic analysis theory of acquisition, induction, and representation of knowledge.' *Psychological Review 104,* 211–240.
Landauer T K, Laham D, Rehder R & Schreiner M E (1997). 'How well can passage meaning be derived without using word order? A comparison of latent semantic analysis and humans.' In *Proceedings of the 19th Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Erlbaum. 412–417.
McNamara D, Landauer T, Dennis S & Kintsch W (eds.) (2005). *LSA: a road to meaning*. Mahwah, NJ: Erlbaum. passim.
Quesada J (2005a). 'Creating your own LSA space.' In McNamara *et al.* (eds.).
Quesada J (2005b). 'Understanding representation in problem solving, judgment and decision making as the creation of multidimensional space.' In McNamara *et al.* (eds.).
Rehder B, Schreiner M, Laham D, Wolfe M, Landauer T & Kintsch W (1998). 'Using Latent Semantic Analysis to assess knowledge: Some technical considerations.' *Discourse Processes 25,* 337–354.
Wiemer-Hastings P, Allbritton D, Efron J & Arnott E (2004). 'RMT: A dialog-based research methods tutor with or without a head.' In *Proceedings of the ITS2004 Seventh International Conference*. Berlin: Springer. 614–623.
Wiemer-Hastings P, Wiemer-Hastings K & Graesser A (1999). 'How latent is latent semantic analysis?' In *Proceedings of the Sixteenth International Joint Congress on Artificial Intelligence*. San Francisco: Morgan Kaufmann. 932–937.
Wiemer-Hastings P & Zipitria I (2001). 'Rules for syntax, vectors for semantics.' In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Erlbaum. 1140–1145.
Wittgenstein L (1958). *Philosophical investigations* (2nd edn.). Anscombe G E M (trans.). Oxford: Blackwell.

Laterality Across the World's Languages

J B Hellige and M M Adamson, University of Southern California, Los Angeles, California, USA

The left cerebral hemisphere is superior to the right for many aspects of speech processing and reading. This overall left-hemisphere dominance is ubiquitous across the world's languages. However, as laterality research has expanded to include a wider variety of languages, we have learned that subtle characteristics of hemispheric asymmetry that were once treated as universal may, instead, depend on specific characteristics of a language. This seems to be particularly true for what has been termed the qualitative nature of processing printed material. With this in mind, this review highlights both the overall consistency and the variability in laterality across the world's languages. Although the review is, of necessity, illustrative rather than exhaustive, it will serve to demonstrate the importance of cross-linguistic laterality studies and to suggest factors that contribute to variability. After a brief review of laterality for identifying speech sounds, we provide a somewhat more extensive review of laterality for identifying printed words and nonwords. In each case, studies using English are reviewed first to provide a context in which to understand studies with other languages.

Processing Speech Sounds

There is a good deal of converging evidence of left-hemisphere superiority for most aspects of speech processing. Though most of the evidence comes from studies of English, laterality for speech processing has been examined across a sufficient variety of languages to demonstrate the generality of left-hemisphere superiority and to suggest a few critical factors that underlie such laterality.

English

Studies of patients with unilateral brain damage, behavioral laterality studies, functional brain imaging studies, and studies using a variety of other techniques all indicate left-hemisphere superiority for identifying acoustically presented words and nonwords. For example, in a behavioral technique known as dichotic listening, two similar but different sounds (e.g., */ba/* and */da/* or */bad/* and */dad/*) are presented simultaneously, one to each ear. Because input pathways from each ear project predominantly to the contralateral cerebral cortex, better identification of the stimuli from one ear is taken to reflect superior processing in the contralateral cerebral hemisphere. For English speakers processing English speech sounds, there are remarkably consistent indications of a right-ear (that is, left-hemisphere) advantage for most people.

Right-ear advantages have been reported for nonwords as well as for words and for a variety of speech sounds including stop consonants, nasals, and vowels (for review, see Hellige, 1993). Recent neuroimaging studies indicate bilateral activation in the region of the superior temporal gyrus, though the amount of activation appears asymmetric. Activation of additional regions depends on the specific task being performed (for review, see Hickok and Poeppel, 2004).

Though there is consistent evidence for a left-hemisphere advantage in identifying which speech sounds have been presented, there are strong indications of a right-hemisphere advantage for identifying such things as the emotional tone conveyed by a word or short sequence of words or for other forms of speech prosody. In fact, within the same dichotic listening study, simultaneous right-ear (left-hemisphere) and left-ear (right-hemisphere) advantages have been shown for identifying word content and emotional tone, respectively (Bryden and MacRae, 1989; Grimshaw *et al*., 2003).

An important area for additional investigation concerns the extent to which effects like those just noted are related to language or to the perception of emotion, per se, or whether they are related to more general hemispheric asymmetry for processing different aspects of acoustic signals. For example, information about the identification of many spoken phonemes in English is conveyed by rapid changes in the acoustic signal over brief periods of time, a type of acoustic change for which the left hemisphere is hypothesized to be well-adapted, whether the stimuli are linguistic or not. In contrast, information about emotional tone of voice is conveyed by much slower modulations of the acoustic signal over longer periods of time, a type of acoustic change for which the right hemisphere is hypothesized to be well adapted (e.g., Bryden and MacRae, 1989; Grimshaw *et al*., 2003; Hellige, 1993).

Other Languages

Laterality studies using a variety of the same techniques that have been used with English demonstrate similar effects across other languages. This includes a variety of European languages as well as languages that, unlike English, use pitch for lexical purposes (e.g., Mandarin Chinese, Thai, and Vietnamese). It is interesting that, in those languages for which phonemic and lexical discriminations are made on the basis of pitch, the left-hemisphere advantage extends to pitch discriminations presented in the context of linguistic discrimination. This stands in contrast to studies of native English speakers in which there is no consistent left-hemisphere dominance (and even some indication of right-hemisphere dominance) for pitch discrimination. Such results do not necessarily argue against the notion that some of the hemispheric asymmetry for speech perception is related to a more general hemispheric asymmetry for processing different types of acoustic information. However, they indicate that, in addition to this, laterality for a specific acoustic property also depends on that property's relevance for making linguistic discrimination in the language being tested.

Processing Printed Words and Nonwords

The left hemisphere is superior to the right for most aspects of processing written or printed words and nonwords. As is the case for speech processing, there are several types of converging evidence, including studies of patients with unilateral brain damage, behavioral laterality studies, and studies of localized brain activity. The greatest amount of evidence again comes from studies of English, though there are now at least some studies in a great many other languages. Though left-hemisphere superiority for certain aspects of processing is widespread, there are interesting exceptions related to specific characteristics of individual languages. As illustrated by the following brief review, variability across languages is especially noteworthy with respect to what have been termed qualitative aspects of processing.

English

A brief review of laterality studies using English provides a context in which to consider laterality across other languages. A great many studies of patients with unilateral brain damage indicate that processing printed words and nonwords is more likely and more severe when the left hemisphere is damaged than when it is not. Furthermore, neuroimaging studies of brain-damaged patients as well as neurologically normal individuals provide evidence of greater left-hemisphere activation during processing of printed words and nonwords in English (e.g., Pugh *et al*., 1997, 2001; Cohen *et al*., 2000; Polk *et al*., 2002; Mechelli *et al*., 2003; Price *et al*., 2003). However, specific cerebral activation patterns differ for various components of reading and may include semantic, phonological, and orthographic networks (Pugh *et al*., 1997, 2001). For instance, silent reading of words, nonwords, letter strings, and psuedowords in English leads to greater activation in left hemisphere, as compared to in the right hemisphere, mainly in the left inferior frontal gyrus, left anterior insula, frontal operculum, left temporoparietal junction, left inferior occipital gryus, and posterior portion of left inferior

temporal area (e.g., Rumsey *et al.*, 1997; Cohen *et al.*, 2000; McCandliss *et al.*, 2003). In addition, nouns and object names in English activate the posterior regions of occipitotemporal and occipital areas, whereas verbs and action names activate the prefrontal and frontal temporal regions (Pulvermuller *et al.*, 1999). Furthermore, the pattern and strength of activation in these areas may depend on the particular properties of different languages that share the same alphabet system. For example, reading of words and nonwords in Italian, an orthographically shallow language compared to English, leads to greater activation in the left superior temporal region as compared to the homologous region in the right hemisphere, whereas reading in English shows greater activation in left frontal and posterior inferior temporal regions (Paulesu *et al.*, 2000).

Converging evidence of functional left-hemisphere superiority also comes from a variety of behavioral studies, including studies that flash visual material very briefly to the left or right of an observer's point of eye fixation. This technique is based on the fact that each visual half-field projects directly to the contralateral brain hemisphere, so that under appropriate conditions, an advantage for one visual field or the other is a reflection of an advantage for one hemisphere or the other. A great many studies in English demonstrate a right visual field (RVF) or left-hemisphere advantage for such things as identifying printed words and nonwords (for review, see Hellige, 1993) or for indicating whether or not a string of printed letters spells a word (a lexical decision task; for examples, see Iacobani and Zaidel, 1996).

In addition to demonstrations of left-hemisphere superiority, there is evidence that printed material may be processed in qualitatively different ways by the left and right brain hemispheres. For example, it has been suggested that the right hemisphere tends to process the letters within a word or even within a pronounceable nonword in a relatively slow, sequential manner, whereas the left hemisphere tends to process the letters within the string in a more holistic or parallel manner. For example, some patients with left-hemisphere injury, whose reading is thought to depend more on the intact right hemisphere, read words laboriously, in a letter-by-letter manner (e.g., Reuter-Lorenz and Baynes, 1992). In addition, in a variety of visual half-field experiments that require word naming or lexical decision, there are word-length effects when stimuli are projected to the left visual field (LVF) or right hemisphere but not when stimuli are projected to the RVF or left hemisphere. That is, in LVF/right hemisphere trials, reaction time and errors increase as word length increases, consistent with the suggestion of slower, serial processing. Interestingly, when words are presented in unusual formats (e.g., the letters are arranged vertically, the letters are misaligned or of mixed case), word-length effects of the same magnitude have been reported for both LVF/right hemisphere and RVF/left hemisphere trials (Young and Ellis, 1985; Lindell *et al.*, 2002). Thus, the ability of the left hemisphere to process printed words in a more parallel manner may be restricted to stimulus arrangements that are typical of English.

Converging evidence for qualitative differences on LVF and RVF trials also comes from studies that required observers to identify consonant-vowel-consonant (CVC) nonword letter sequences (e.g., 'dag'). Across many such studies, there is a robust RVF/left hemisphere advantage for this task. In addition, distinctively different error patterns are obtained on LVF and RVF trials. On LVF trials, the third letter is missed more often than the first letter, just as one would expect if the letters were being processed in a relatively slow, sequential manner. On RVF trials, this difference between first and third letter errors is reduced and may even be eliminated if the letters are arranged in the typical horizontal manner. It has been suggested that this reflects the special ability of the left hemisphere to treat the CVC stimulus as a single pronounceable unit rather than as a simple collection of individual letters (for examples and review, see Hellige and Marks, 2001; Marks and Hellige, 1999, 2003). Thus, examination of these qualitative error patterns may shed light on the manner in which processing depends on characteristics that differ across the world's languages. In fact, as the subsequent review demonstrates, qualitative aspects of processing seem to be quite a bit more variable than the more prevalent quantitative advantage for the left hemisphere.

Much research with European languages produces laterality patterns very similar to the patterns reported for English and, consequently, will not be reviewed here. Instead, we focus on selected languages whose laterality patterns differ in interesting ways from those found with English. To help set a context for understanding the laterality effects, each of the following language-specific sections begins with a brief overview of the characteristics of printed text in that language and is followed by a summary of what is known about laterality for processing that text. Following these sections, we suggest some general factors that may be related to variability across languages in both quantitative and qualitative aspects of laterality.

The proto-Semitic system that gave rise to Hebrew, Arabic, Farsi, Urdu, and Sanskrit consisted mainly of consonants; the markings of vowels and their pronunciations was made optional. These scripts continue to

have several common features, including the fact that the modern versions of each typically omit vowel markings and the fact that all except Sanskrit are scanned from right to left. Of course, each of these languages also has its own unique morphology as well as other characteristics that may influence laterality. With these things in mind, the next several sections review laterality for Hebrew, Arabic, Farsi, and Urdu.

Hebrew

Written Hebrew words are produced when semantic roots are embedded with vocalic indications and syllabic prefixes or suffixes. Written Hebrew consists of a string of consonants, in the sense that there are no separate characters to designate vowels. Instead, four of the consonantal letters are also used to designate vowels, creating a good deal of ambiguity. A 'pointed' form of text uses diacritical marks to identify correct vowel pronunciation and location. This form has perfect grapheme-to-phoneme correspondence so that words can be easily pronounced. The typical modern form of Hebrew text is 'unpointed,' omitting the diacritical marks, producing a deep orthography and many homographs (words that are spelled the same but pronounced differently). This is generally thought to lead to a lexical reading strategy, in which an entire word is processed before phonological information is extracted postlexically (Bentin *et al.*, 1984; Share, 2003). Just as in English, processing of printed Hebrew text is disrupted more by left-hemisphere damage than by right-hemisphere damage. At the same time, more subtle differences in dyslexic symptoms are almost certainly related to specific characteristics of Hebrew. For example, studies of surface dyslexia in Hebrew speakers show that the patients are able to read nonwords and other phonological transformations when the material is presented in pointed form but are unable to read unpointed Hebrew (Birnboim, 1995). Whereas in English surface dyslexics have difficulty with irregular (exception) words (Marshall and Newcombe, 1973; Patterson, 1981; Coltheart *et al.*, 1983) most of the errors made in Hebrew are attributable to misreading of the vowels, leading to a very selective type of peripheral dyslexia. The specific function impaired in these patients is the ability to locate the position of a vowel within a word while letter identification itself is intact (Friedman and Gvion, 2001).

In behavioral studies of Hebrew, there is a robust RVF/left hemisphere advantage for naming printed words and nonwords as well as for making lexical decisions (Faust *et al.*, 1993; Babkoff *et al.*, 1997; Lavidor *et al.*, 2002). Thus, quantitative aspects of laterality are very similar to those found with English. However, qualitative aspects of laterality are sometimes quite different. For example, experiments investigating word-length effects in Hebrew have produced mixed results. Similar to English, some studies show word length has an effect only on LVF/right hemisphere trials in Hebrew during lateralized lexical decision tasks (Lavidor *et al.*, 2002). However, other studies, using similar methodology, show length effects in both visual fields (Koriat, 1985; Lavidor *et al.*, 2001), and studies of other types indicate the opposite result from English (Nazir *et al.*, 2004). In addition, in studies that required identification of CVC nonword stimuli, the visual field difference in qualitative error pattern was opposite that reported for similar studies English, at least when the letters were presented vertically. That is, there were more third-letter errors than first-letter errors on RVF/left hemisphere trials, and this difference was reduced on LVF/right hemisphere trials (Eviatar, 1999). It is important to keep in mind the fact that even though this qualitative aspect of laterality was reversed for the two languages, there were equally robust RVF/left hemisphere accuracy advantages in both. In considering the qualitative differences between Hebrew and English, Eviatar (1999) suggests that relevant factors may include the difference in scanning direction as well as the possibility that the morphology of unvoweled (unpointed) Hebrew (which was used in her study) is such that the entire stimulus may have to be scanned because letters of the root are dispersed across entire words. From this perspective it is interesting that there was no visual field difference in qualitative error patterns when the Hebrew nonwords were presented with the letters arranged in the more typical horizontal manner.

Arabic

In Arabic, Farsi and Urdu letters are connected in script-like fashion to form words or pronounceable nonwords. This means that each letter may be represented by two to four distinct allographs, depending on its position in word-initial, word-medial, word-final, or isolation. For Arabic, 22 of 28 letters in the alphabet have four shapes each (word initial, medial, final, and when they come after an unconnecting letter). Six other letters have two shapes each (Ibrahim *et al.*, 2002). Arabic has two forms: literary Arabic, which is taught in schools as the standard script for reading and writing, and modern Arabic, which is the native language of Arabs. Although the two forms of Arabic share orthographic properties, they are phonologically and syntactically different. More specifically, certain vowels exist in modern Arabic but are considered illegal in literary Arabic. Other

differences include usage of different word order constraints in sentence structure, employing different plural markings and insertion rules for function words (Bentin and Ibrahim, 1996). Arabic orthography, similar to Hebrew, has two forms, one in which the vowels are marked, and another in which they are not marked. As in Hebrew, if all the vowels are marked, above or below the letters in a word, literary Arabic is phonologically transparent. If the vowel marks are not included, the orthography becomes opaque (Bentin and Ibrahim, 1996). However, unlike Hebrew, the four letters that designate long vowels in Arabic do not also function as consonants, and inclusion of them in a word helps the pronunciation. Modern Arabic does not include the vowel markings, but the letters designating the vowels are never omitted, and their locations within a word indicate the location of the vowels. The number and location of the dots on a similar-looking letter can change the grapheme it represents. For example, the Arabic letters representing */t/* and */n/* can become */th/* and */b/* by changing the number and location of the dots.

Although there is not much research on language disorders involving Arabic speakers, a case study of an Arabic/French bilingual reports symptoms of deep dyslexia in both languages after damage to the left perisylvian region. However, the type of errors made was different in each language. For instance, more semantic errors occurred during picture naming and reading aloud in French than in Arabic. In contrast, there were more translation errors during reading in Arabic (e.g., the Arabic word for 'glasses' was read as the French word for 'glasses'; Beland and Mimouni, 2001).

Laterality research involving the normal Arabic-speaking population is also quite sparse. Studies of letter recognition and letter matching show that Arabic orthography is more difficult than Hebrew orthography, and the pattern of errors is such that Arabic speakers have most difficulty when Arabic letters are presented to the LVF/right hemisphere (Eviatar *et al*., 2004). In a study using vertical presentation of Arabic CVC stimuli, Arabic–Hebrew bilinguals showed a robust RVF/left hemisphere advantage and a visual field difference in error pattern that was similar to that shown when Hebrew speakers identified Hebrew CVC stimuli. That is, the extent to which the third letter was missed more often than the first letter was larger on RVF/left hemisphere trials than on LVF/right hemisphere trials, just the opposite of English. In addition, averaged across both visual fields, the difference between third-letter and first-letter errors was larger for Arabic than for English or Hebrew. Perhaps the greater orthographic complexity of Arabic combined with the unfamiliar vertical orientation led to more sequential processing being used in both visual fields. Further studies would be useful to examine these possibilities and others.

Farsi

Written Farsi adopted the Nastaliq script in about 1500 A.D. (Hanaway and Spooner, 1995). Nastaliq contains the Arabic characteristics of dots identifying different graphemes and of joining different letters to form a word. Farsi has six vowels, three of which are represented by letters in the word. The other three are represented by vowel markings placed above (two) and below (one) the letters. These marks are used only for beginning readers (similar to literary Arabic and pointed Hebrew), as adults read text without these markings. However, the three letters designating the vowels are never omitted. Whereas stimuli with a fixed letter marking a vowel are linguistically transparent, stimuli with no vowel markings are linguistically opaque. In Farsi, the transparent words are read by using a phonological, nonlexical route, and the opaque words are read by using a lexical route (Baluch and Besner, 1991, 2001).

In one of the only laterality studies using Farsi, Melamed and Zaidel (1993) presented Farsi and English to Farsi–English bilinguals during naming and lexical decision tasks in a visual half-field paradigm. The results revealed a language by visual field interaction, showing the typical RVF/left hemisphere advantage in accuracy and reaction time during naming and lexical decision tasks in English, but no such advantage during naming and lexical decision tasks in Farsi. The authors suggest that the absence of a RVF/LH advantage may indicate greater involvement of the right hemisphere in processing written Farsi, perhaps because of its visual or orthographic complexity or because it is typically scanned from right to left. In any case, the absence of an RVF/LH advantage is sufficiently unusual among the world's alphabetic written languages to warrant replication and extension.

Urdu

Like Arabic and Farsi, Urdu is scanned from right to left, and an important difference from English stems from its orthographic property that clusters of letters together do not make a word (or even a pronounceable nonword) unless they are connected in script-like fashion. Unlike English, there are no distinctions between upper and lower cases in Urdu, but a letter's image may change when it is joined in script with other letters. Designated vowels (a, e, i, o, and u) can be in any position within a word or pronounceable nonword. In some literary and religious texts, vowel markings are used to accentuate spoken vowels. In

everyday Urdu, like Hebrew, Arabic, and Farsi, vowel markings are not used. Urdu self-teaching books designate each unit of sound to one letter. Native speakers employ this letter-by-letter reading strategy as an initial learning technique when they learn Urdu formally. Once basic Urdu is learned, and the patterns of words are understood, this letter-by-letter strategy is abandoned for a more global word-reading technique.

In a series of visual half-field studies, Urdu–English bilinguals were asked to identify sequences of three Urdu letters that were either joined together so as to make Urdu words or pronounceable nonwords or separated so as to be spontaneously processed as individual letters (M. M. Adamson and J. B. Hellige, unpublished data). For both joined and separated words and non-words, there was a RVF/left hemisphere advantage similar to that reported for many other languages. Overall, the third letter was missed more often than the first letter, and this difference was larger for separated letter sequences than for joined letter sequences. This indicates that Urdu letters are processed in a slower, more sequential manner when they are separated than when they are joined to make a word or pronounceable nonword. Visual field differences in the pattern of errors were similar to those found for Hebrew and Arabic and opposite that found for English. Specifically, the tendency for there to be more third-letter errors than first-letter errors was larger on RVF/left hemisphere trials than on LVF/right hemisphere trials. The laterality patterns were identical for words and nonwords, indicating that the effects occur at a prelexical stage of processing (for similar arguments regarding effects in English, see Lindell *et al*., 2002). The laterality patterns were also similar for joined and separated versions of the stimuli, which may not be surprising in view of the fact that both versions are made up of the same letters, written in the same direction and in the same sequence.

Japanese

The Japanese language system consists of two different types of script: *Kanji*, a logographic form borrowed from Chinese characters, and varieties of the syllabic form of *Kana*, mainly, *Hiragana* and *Katakana*. A Japanese sentence is written by combining all three of these scripts to form different classes of words (Wydell, 2003). Japanese *Kana* has a direct one-to-one relationship between its orthographic representation and its pronunciation. Because of this direct link between phonology and orthography, children learn the *Kana* script very quickly in school (Wydell, 2003). Japanese *Kanji*, however, is taught at junior or high school level. The most striking difference between *Kanji* and *Kana* is that *Kana* depends on the phonological system for its correct pronunciation, and *Kanji* is more dependent on the visual spatial system for correct processing. This is because each *Kanji* character is made up of a morphological element that has no phoneme component or correspondence; that is, a *Kanji* character does not map directly onto any particular phoneme (Wydell, 1998).

Case studies of brain-damaged patients describe the symptoms of phonological alexia in *Kana* and *Kanji* as the inability to read nonwords in both scripts (Patterson *et al*., 1996; Sasanuma *et al*., 1996), though it has been suggested that the two scripts use somewhat different neural networks. It is often extremely difficult to compare neuroimaging studies across languages because of differences in various aspects of protocol, including the choice of specific experimental and baseline task demands. Despite these difficulties, a number of studies now indicate that processing of both *Kana* and *Kanji* generally activate the same cortical areas as each other and as are activated in similar studies using English. At the same time, and despite the activation of overlapping cortical regions, it appears that processing *Kanji* involves heavier visual orthographic demands and uses a lexical-semantic system through a more ventral route, involving the inferior occipital, fusiform, and posterior inferior temporal gyri, whereas *Kana* is processed in a more phonetic manner through a more dorsal route, involving temporoparital areas (for elaboration and review, see Thuy *et al*., 2004).

Behavioral studies using the visual half-field technique have sometimes reported a RVF/left hemisphere advantage for identifying *Kana* words and nonwords, but a LVF/right hemisphere advantage for identifying *Kanji* words and nonwords (Hatta, 1977; Sasanama *et al*., 1977), leading some to conclude that the left and right hemispheres are superior for processing *Kana* and *Kanji*, respectively (e.g., Coltheart, 1980). It has been suggested that it is the visual complexity of *Kanji* stimuli that leads to more dependence on right-hemisphere processing, with right-hemisphere dominance for visuospatial processing outweighing any left-hemisphere superiority for other stages of processing involved in stimulus identification. In contrast to English, there may not be any laterality effects for lexical decision tasks for *Kanji* (Hatta, 1981), though there is a RVF/LH advantage for lexical decision tasks for *Katakana* and *Hiragana* (Yoshizaki, 2001). A RVF/LH advantage for *Kanji* emerges, however, if the task requires a more semantic judgment (Hatta, 1981).

In a visual half-field field study that required native speakers of Japanese to identify three-letter *Kana* nonwords presented vertically, Hellige and Yamauchi

(1999) found a robust RVF/left hemisphere advantage. Although there were more third-letter errors than first-letter errors, consistent with sequential processing of the letters, this was equally true for both LVF and RVF trials. Hellige and Yamauchi suggest that this may stem from the way in which *Kana* maps on to pronunciation. Specifically, a word or nonword made up of *Kana* characters is pronounced exactly as it is spelled: letter by letter. That is, there is no emergent or unitary pronunciation, as there is in English (D-A-G is pronounced *daeg*), and consequently, less reason for the specialized language mechanisms of the left hemisphere to process the letter sequence in a more holistic way. In addition, *Kana* is most typically scanned from top to bottom, which is quite different from the languages discussed earlier, and this may contribute to the lack of visual field differences in qualitative error patterns.

Chinese

Chinese is a nonalphabetic language with deep orthography (DeFrancis, 1989). Chinese builds its single graphic units from independent strokes, making compound units by incorporating the graphic units in different ways; mainly, left and right, top and bottom, and inside and outside (Perfetti *et al.*, 2002). The main phonetic unit of Chinese Mandarin is a syllable that is further divided into two segments, initial and final. The main script unit is a monosyllabic character representing units of meaning (morphemes). About 90% of Chinese characters (referred to as morphemic) are compound characters having both a semantic and a sound component (Zhu, 1987). Therefore, only 10% of Chinese characters (referred to as pictorial) are exclusively pictographic. The number of strokes in a character can change its phonological or semantic properties (Ho, 2003). Unlike English, recent studies show that reading Chinese involves orthographic-semantic-phonological access at both the word and character levels (Zhang and Perfetti, 1993; Kuo *et al.*, 2001, 2003; Bates *et al.*, 2003). In addition, each character has a phonetic value, but its tone depends on the location within a word, which can then change the meaning (Ho, 2003). In contrast to Chinese, *pinyin* is a completely syllabic script considered to be alphabetic and was created in the 1970s to introduce English words in the mainstream Chinese language. The distinction between *pinyin* and Chinese is similar to the distinction between Japanese *Kana* and *Kanji* and provides interesting opportunities for comparison that are only beginning to be exploited.

Neuroimaging studies indicate that both Chinese and *pinyin* characters activate the same general cortical areas as each other and as are activated in other languages. However, relative activation in some regions differs for these two forms. It has been suggested that activity in some of these regions (e.g., the inferior parietal area for *pinyin* and the fusiform gyrus for Chinese characters) may be associated with reading via assembled phonology (*pinyin*) versus addressed whole-script (Chinese character) procedures (Chen *et al.*, 2002). A variety of additional neuroimaging studies requiring individuals to perform a number of different tasks with Chinese characters is generally consistent with the conclusion that the left inferior frontal gyrus, the left temporoparietal region, and the left temporal-occipital junction are involved in orthographic-to-phonological translation for Chinese, as well as for a number of other languages (e.g., Lee *et al.*, 2004). At the same time, there is some indication that a region within the left middle frontal cortex may be activated in phonological processing of written Chinese but not in processing of most alphabetic languages and that the right hemisphere is heavily involved in reading aloud of Chinese logographs, which also differs from alphabetic languages (e.g., Tan *et al.*, 2001). And the different activation patters found for nouns and verbs in English does not extend to Chinese (e.g., Li *et al.*, 2004). Thus, like alphabetic languages, greater left-hemisphere activation in a number of areas is generally obtained during processing characters and words in Chinese. But, somewhat like Japanese *Kanji*, the subtle differences in neuronal activity echo the differences embedded in the linguistic and visually complex properties of Chinese that make it different from alphabetic languages (for additional neuroimaging studies using a variety of tasks, see Chee *et al.*, 2000; Gandour *et al.*, 2003; Kuo *et al.*, 2003, 2004).

Similar to Japanese *Kanji*, visual half-field studies of Chinese involving normal Chinese speakers reveal a LVF/right hemisphere advantage in word naming and no visual field advantage in a lexical decision task (Tzeng *et al.*, 1979), with this being the case for both pictorial and morphemic Chinese characters (Tzeng *et al.*, 1978; but see Rastatter *et al.*, 1989). As with Japanese *Kanji*, however, a RVF/left hemisphere advantage is found for Chinese when deeper phonological or semantic processing tasks are used (e.g., Leong *et al.*, 1985). In contrast to English, there may not be any word length by visual field advantage for processing Chinese characters, though such an interaction has been reported when the Chinese characters to be identified were embedded in foreign words of different lengths (e.g., Fang, 1994, 1997, 2003) Thus, some aspects of processing Chinese characters and words are qualitatively different from processing English and other languages that more readily show word length by visual field interactions.

General Observations

The foregoing review indicates that there is widespread left-hemisphere superiority for processing printed words and nonwords. This is true for many of the world's languages, and for at least certain aspects of orthographic to phonological mapping, left-hemisphere dominance seems nearly universal. At the same time, printed languages differ in many ways, and there is clear evidence that those differences affect certain aspects of laterality. As examples, consider possible effects related to visual complexity and scanning direction.

One way in which languages differ is in terms of the level of visual complexity represented in printed characters. The right hemisphere is hypothesized to be superior to the left in processing complex visual information outside the language domain. Thus, it would not be surprising to find that the right hemisphere makes a greater contribution to processing texts as they become more visually complex. From this perspective, it may be understandable that the RVF/left hemisphere advantage so typical of letter and word identification in English does not extend to character and word identification in visually complex texts like Japanese *Kanji* and Chinese.

Languages also differ in the direction in which printed material is typically scanned, and this may well have consequences for such things as behavioral measures of laterality. To be sure, RVF/left hemisphere advantages have been reported for left-to-right languages (like English), for right-to-left languages (like Hebrew, Arabic, and Urdu), and for top-to-bottom languages (like Japanese *Kana*). Thus, an overall left-hemisphere advantage for such things as word naming seems sufficiently robust to overcome any effects of scanning direction. However, more subtle qualitative performance differences between LVF/right hemisphere and RVF/left hemisphere trials are more variable across languages and are more likely to be influenced by such things as scanning direction. Consider, for example, the finding reviewed earlier that visual field differences in qualitative error patterns obtained during CVC identification are opposite for English (a left-to-right language) and Hebrew, Arabic, and Urdu (all right-to-left language), and that the qualitative visual field difference is not present at all for Japanese *Kana* (primarily a top-to-bottom language). Of course, we cannot be certain that scanning direction is the only relevant factor in determining these qualitative differences, but the association is nonetheless interesting. In fact, such effects of scanning direction are consistent with a recent model of letter-position encoding that specifies how the habitual scanning direction of a language can produce asymmetric activation patterns across a string of letters and thereby lead to visual field differences in word length and error pattern (Whitney, 2001; Whitney and Lavidor, 2004).

Concluding Comments

In reviews of laterality for language-related processing, it is common to focus on studies that ask native speakers of English to engage in a variety of tasks using English. It is often tempting to generalize the results for English and reach conclusions about laterality for language more generally. Though much has been learned in this way and many of the results obtained with English do, in fact, generalize to many other languages, this review indicates that there are important exceptions, especially in qualitative aspects of processing. Understanding both the generality and the variability in laterality across the world's languages is essential to refining our theories about the brain mechanisms that underlie our ability to understand the spoken and the written word.

See also: Anatomical Asymmetries versus Variability of Language Areas of the Brain; Language in the Nondominant Hemisphere.

Bibliography

Babkoff H, Faust M & Lavidor M (1997). 'Lexical decision, visual hemifield and angle of orientation.' *Neuropsychologia 35,* 487–495.

Baluch B & Besner D (1991). 'Visual word recognition: evidence for strategic control of lexical and nonlexical routes in oral reading.' *Journal of Experimental Psychology: Learning, Memory and Cognition 17,* 644–652.

Baluch B & Besner D (2001). 'Basic processes in reading: semantics affects speeded naming high-frequency words in an alphabetic script.' *Canadian Journal of Experimental Psychology 55,* 63–69.

Bates E, D'Amico S, Jacobsen T, Szskely A, Andonova E, Devescovi A, Herron D, Lu C C, Pechmann T, Pleh C, Wicha N, Federmeier K, Gerdjikova I, Gutierrez G, Hung D, Jeanne H, Iyer G, Kohnert K, Mehotcheva T, Orozco-Figueroa A, Tzeng A & Tzeng O (2003). 'Timed picture naming in seven languages.' *Psychonomic Bulletin Review 10,* 344–380.

Beland R & Mimouni Z (2001). 'Deep dyslexia in the two languages of an Arabic/French bilingual patient.' *Cognition 82,* 77–126.

Bentin S, Bargai N & Katz L (1984). 'Orthographic and phonemic coding for lexical access: evidence from Hebrew.' *Journal of Experimental Psychology: Learning, Memory, and Cognition 10,* 353–368.

Bentin S & Ibrahim R (1996). 'New evidence for phonological processing during visual word recognition: the case of Arabic.' *Journal of Experimental Psychology: Learning, Memory, and Cognition 22,* 309–323.

Birnboim S (1995). 'Acquired surface dyslexia: the evidence form Hebrew.' *Applied Psycholinguistics 16,* 83–102.

Bryden M P & MacRae L (1989). 'Dichotic laterality effects obtained with emotional words.' *Neuropsychiatry, Neuropsychology, Behavioral Neurology 1,* 171–176.

Chee M W L, Weekes B, Lee K M, Soon C S, Schreiber A, Hoon J J & Chee M (2000). 'Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: evidence from fMRI.' *NeuroImage 12,* 392–403.

Chen Y, Fu S, Iversen S D, Smith S M & Matthews P M (2002). 'Testing for dual brain processing routes in reading: a direct contrast of Chinese character and Pinyin reading using fMRI.' *Journal of Cognitive Neuroscience 14,* 1088–1098.

Cohen L, Dehaene S, Chochon F, Lehericy S & Naccache L (2000). 'Language and calculation within the parietal lobe: a combined cognitive, anatomical and fMRI study.' *Neuropsychologia 38,* 1426–1440.

Coltheart M (1980). 'Deep dyslexia: a right hemisphere hypothesis.' In Coltheart M, Patterson K & Marshall J C (eds.) *Deep dyslexia.* London: Routledge and Kegan Paul.

Coltheart M, Masterson J, Byng S, Prior M & Riddoch J (1983). 'Surface dyslexia.' *Quarterly Journal of Experimental Psychology: Human Experimental Psychology 35,* 469–495.

DeFrancis J (1989). *Visual speech: the diverse oneness of writing systems.* Honolulu: University of Hawaii Press.

Eviatar Z (1999). Cross-language tests of hemispheric strategies in reading nonwords. *Neuropsychology 13,* 498–515.

Eviatar Z, Ibrahim R & Ganayim D (2004). 'Orthography and the hemispheres: visual and linguistic aspects of letter processing.' *Neuropsychology 18,* 174–184.

Fang S P (1994). 'English word length effects and the Chinese character-word difference: truth or myth?' *Chinese Journal of Psychology 36,* 59–80.

Fang S P (1997). 'Morphological properties and the Chinese character-word difference in laterally patterns.' *Journal of Experimental Psychology: Human Perception and Performance 23,* 1439–1453.

Fang S P (2003). 'Are there differential word length effects in the two visual fields?' *Brain and Language 85,* 467–485.

Faust M, Kravetz S & Babkoff H (1993). 'Hemispheric specialization of reading habits: evidence from lexical decision research with Hebrew words and sentences.' *Brain and Language 44,* 254–263.

Friedman N & Gvion A (2001). 'Letter position dyslexia.' *Cognitive Neuropsychology 18,* 673–696.

Gandour J, Wong D, Dzemidzic M, Lowe M, Tong Y & Li X (2003). 'A cross-linguistic fMRI study of perception of intonation and emotion in Chinese.' *Human Brain Mapping 18,* 149–157.

Grimshaw G M, Kwasny K M, Covell E & Johnson R A (2003). 'The dynamic nature of language lateralization: effects of lexical and prosodic factors.' *Neuropsychologia 41,* 1008–1019.

Hanaway W L & Spooner B (1995). *Reading Nastaliq: Persian and Urdu hands, from 1500 to the present.* Costa Mesa: Mazda Publishers.

Hatta T (1977). 'Recognition of Japanese Kanji in the left and right visual fields.' *Neurospsychologia 15,* 685–688.

Hatta T (1981). 'Different stages of Kanji processing and their relations to functional hemisphere asymmetries.' *Japanese Psychological Research 23,* 27–36.

Hellige J B (1993). *Hemispheric asymmetry: what's right and what's left.* Cambridge, MA: Harvard University Press.

Hellige J B & Marks N L (2001). 'Are hemispheric strategy differences independent of level of performance?' *Neuropsychology 15,* 380–395.

Hellige J B & Yamauchi M (1999). 'Quantitative and qualitative hemispheric asymmetry for processing Japanese Kana.' *Brain and Cognition 40,* 453–463.

Hickok G & Poeppel D (2004). 'Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language.' *Cognition 92,* 67–99.

Ho C S (2003). 'Reading acquisition and developmental dyslexia in Chinese: a cognitive perspective.' In Goulandris N (ed.) *Dyslexia in different languages: cross-linguistic comparisons.* London: Whurr Publishers.

Iacoboni M & Zaidel E (1996). 'Hemispheric independence in word recognition: evidence from unilateral and bilateral presentations.' *Brain and Language 53,* 121–140.

Ibrahim R, Eviatar Z & Aharon-Peretz J (2002). 'The characteristics of Arabic orthography slows its processing.' *Neuropsychology 16,* 322–326.

Koriat A (1985). 'Lateralization effects in pointed and unpointed Hebrew.' *British Journal of Psychology 76,* 161–173.

Kuo W, Yeh T, Duann J, Wu Y, Ho L, Hung D, Tzeng O & Hsieh J (2001). 'A left-lateralized network for reading Chinese words: a 3T fMRI study.' *Neuroreport 12,* 3997–4001.

Kuo W-J, Yeh T-C, Lee C-Y, Lee C-Y, Wu Y-T, Chou C-C, Ho L-T, Hung D L, Tzeng O J L & Hsieh J-C (2003). 'Frequency effects of Chinese character processing in the brain: an event-related fMRI study.' *NeuroImage 18,* 720–730.

Kuo W-J, Yeh T-C, Lee J-R, Chen L-F, Lee P-L, Chen S-S, Ho L-T, Hung D L, Tzeng O J-L & Hsieh J-C (2004). 'Orthographic and phonological processing of Chinese characters: an fMRI study.' *NeuroImage 21,* 1721–1731.

Lavidor M, Babkoff H & Faust M (2001). 'Analysis of standard and non-standard visual format in the two hemispheres.' *Neuropsychologia 39,* 430–439.

Lavidor M, Ellis A W & Pansky A (2002). 'Case alteration and length effect in lateralized word recognition: studies of English and Hebrew.' *Brain and Cognition 50,* 257–271.

Lee C-Y, Tsai J-L, Kuo W-J, Yeh T-C, Wu Y-T, Ho L-T, Hung D L, Tzeng O J L & Hsieh J-C (2004). 'Neuronal correlates of consistency and frequency effects on Chinese character naming: an event-related fMRI study.' *NeuroImage 23,* 1235–1245.

Leong C K, Wong S, Wong A & Hiscock M (1985). 'Differential cerebral involvement in perceiving Chinese characters: levels of processing approach.' *Brain and Language 26,* 131–145.

Li P, Jin Z & Tan L H (2004). 'Neural representations of nouns and verbs in Chinese: an fMRI study.' *NeuroImage 21,* 1533–1541.

Lindell A, Nicholls M E R & Castles A E (2002). 'The effect of word length on hemispheric word recognition: evidence from unilaeral and bilateral-redundant presentations.' *Brain and Cognition 48,* 447–452.

Marks N L & Hellige J B (1999). 'Effects of bilateral stimulation and stimulus redundancy on interhemispheric interaction.' *Neuropsychology 13,* 475–487.

Marks N L & Hellige J B (2003). 'Interhemispheric interaction in bilateral redundancy gain: effects of stimulus format.' *Neuropsychology 17,* 578–593.

Marshall J C & Newcombe F (1973). 'Patterns of paralexia: a psycholinguistic approach.' *Journal of Psycholinguistic Research 2,* 175–199.

McCandliss B D, Cohen L & Dehaene S (2003). 'The visual word form area: expertise for reading in the fusiform gyrus.' *Trends Cognitive Science 7,* 293–299.

Mechelli A, Gorno-Tempini M L & Price C J (2003). 'Neuroimaging studies of word and pseudoword reading: consistencies, inconsistencies, and limitations.' *Journal of Cognitive Neuroscience 15,* 260–271.

Melamed F & Zaidel E (1993). 'Language and task effects on lateralized word recognition.' *Brain and Language 45,* 70–85.

Nazir T A, Ben-Boutayab N, Decoppet N, Deutsch A & Frost R (2004). 'Reading habits, perceptual learning, and recognition of printed words.' *Brain and Language 88,* 294–311.

Patterson K (1981). 'Reading disorders in adults: an information processing approach.' *IPO Annual Progress Report 16,* 96–98.

Patterson K, Suzuki T & Wydell T N (1996). 'Interpreting a case of Japanese phonological alexia: the key is in phonology.' *Cognitive Neuropsychology 13,* 803–822.

Paulesu E, McCrory E, Fazio F, Menoncello L, Brunswick N, Cappa S, Cotelli M, Cossu G, Corte F, Lorusso M, Pesenti S, Gallagher A, Perani D, Price C, Frith C & Frith U (2000). 'A cultural effect on brain function.' *Nature Neuroscience 3,* 91–96.

Perfetti C A, Liu Y & Tan L-H (2002). 'How the mind can meet the brain in reading: a comparative writing systems approach.' In Kao H S R, Leong C-K & Gao D G (eds.) *Cognitive neuroscience studies of the Chinese language.* Hong Kong: Hong Kong University Press.

Polk T A, Stallcup M *et al.* (2002). 'Neural specialization for letter recognition.' *Journal of Cognitive Neuroscience 14,* 145–159.

Price C J, Gorno-Tempini M L, Graham K S, Biggio N, Mechelli A, Patterson K & Nappeney U (2003). 'Normal and pathological reading: converging data from lesion and imaging studies.' *Neuroimage 20,* S30–S41.

Pugh K R, Mencl E E, Jenner A R, Katz L, Frost S J & Lee J R (2001). 'Neurobiological studies of reading and reading disability.' *Journal of Communication Disorders 34,* 479–492.

Pugh K R, Shaywitz B A, Shaywitz S R, Shankweiler D, Katz L, Fletcher J, Constable T, Skudlarski P, Fullbright R, Bronen R & Gore J (1997). 'Predicting reading performance from neuroimaging profiles: the cerebral basis of phonological effects in printed word identification.' *Journal Experimental Psychology Human Perception Performance 23,* 299–318.

Pulvermuller F, Lutzenberger W & Preissl H (1999). 'Nouns and verbs in the intact brain: evidence from event-related potentials and high-frequency cortical responses.' *Cerebral Cortex 9,* 497–506.

Rastatter M, Scukanec G & Grilliot J (1989). 'Hemispheric specialization for processing Chinese characters: some evidence from lexical decision vocal reaction times.' *Perceptual and Motor Skills 69,* 1083–1089.

Reuter-Lorenz P A & Baynes K (1992). 'Modes of lexical access in the callosotomized brain.' *Journal of Cognitive Neuroscience 4,* 155–164.

Rumsey J M, Horwitz B, Donohue B C, Nace K, Maisog J M & Andreason P (1997). 'Phonological and orthographic components of word recognition. A PET-rCBF study.' *Brain 120,* 739–759.

Sasanuma S, Itoh M, Moi K & Kobayashi Y (1977). 'Tachistoscopic recognition of Kana and Kanji words.' *Neuropsychologia 15,* 547–553.

Sasanuma S, Ito H, Patterson K & Ito T (1996). 'Phonological alexia in Japanese: a case study.' *Cognitive Neuropsychology 13,* 823–848.

Share D L (2003). 'Dyslexia in Hebrew.' In Goulandris N (ed.) *Dyslexia in different languages: cross-linguistic comparisons.* London: Whurr Publishers Ltd.

Tan L H, Feng C-M, Fox P T & Gao J-H (2001). 'An fMRI study with written Chinese.' *Neuroreport 12,* 83–88.

Thuy D H, Matsuo K, Nakamura K, Toma K, Oga T, Nakai T, Shibasaki H & Fukuyama H (2004). 'Implicit and explicit processing of kanji and kana words and non-words studied with fMRI.' *NeuroImage 23,* 878–889.

Tzeng O, Hung D L, Cotton B & Wang W S-Y (1979). 'Visual lateralization effect in reading Chinese characters.' *Nature 282,* 499–501.

Tzeng O, Hung D L & Garro L (1978). 'Reading the Chinese characters: an information processing view.' *Journal of Chinese Linguistics 6,* 287–305.

Whitney C (2001). 'How the brain encodes the order of letters in a printed word: the SERIOL model and selective literature review.' *Psychonomic Bulletin and Review 8,* 221–243.

Whitney C & Lavidor M (2004). 'Why word length only matters in the left visual field.' *Neuropsychologia 42,* 1680–1688.

Wydell T N (1998). 'What matters in Kanji word naming: consistency, regularity, or ON/KUN-reading difference?' *Reading and Writing: An Interdisciplinary Journal 10,* 359–373.

Wydell T N (2003). 'Dyslexia in Japanese and the "hypothesis of granularity and transparency."' In Goulandris N

(ed.) *Dyslexia in different languages: cross-linguistic comparisons.* London: Whurr Publishers.
Yoshizaki K (2001). 'Effects of visual familiarity for words on interhemispheric cooperation for lexical processing.' *Cognitive Brain Research 12,* 409–414.
Young A W & Ellis A W (1985). 'Different methods of lexical access for words presented in the left and right visual hemifields.' *Brain and Language 24,* 326–358.
Zhang S & Perfetti C A (1993). 'The tongue-twister effect in reading Chinese.' *Journal of Experimental Psychology: Learning, Memory, and Cognition 19,* 1082–1093.
Zhu X (1987). 'Analysis of cueing function of phonetic components in modern Chinese.' In Yuan X (ed.) *Proceedings of the symposium on the Chinese language and characters.* Beijing: Guang Ming Daily Press.

Latin

J Powell, Royal Holloway, University of London, Egham, Surrey, UK

History and Affiliations

Latin is an Indo-European language of western central Italy, recorded from about 500 B.C. in the immediate area of Rome, related closely to the poorly attested Faliscan dialect and more distantly to Osco-Umbrian and Venetic. Lexical influences from neighboring languages (Italic dialects, the unrelated Etruscan language, and above all Greek) are visible in Latin from the earliest period. With the rise of Rome, Latin spread throughout Italy and into the provinces of the Roman Empire. It became the spoken language in western continental Europe, displacing the indigenous languages; its spoken varieties developed into the modern Romance languages. A more or less standard version of written Latin remained in use in administration, law, education, and the Church, and in due course spread throughout Europe as Medieval Latin. The status of Latin as a common medium of learned communication began to wane only with the rise of English as a world language in modern times.

Varieties

The dialect of Rome became the dominant form of Latin from an early period, with variations from it being regarded as rustic. Written Latin was standardized by the first century B.C. and remained broadly uniform thereafter, concealing changes and diversification in the spoken idiom; nevertheless a range of literary, technical, or colloquial registers can be distinguished in the surviving texts, and the occurrence of nonstandard written forms can provide evidence for spoken developments. The term *Vulgar Latin* is used to refer to any variety that departs significantly from classical norms. There is no good ground for regarding later written forms – Christian, Medieval, Renaissance, or Modern Latin – as essentially different varieties of the language: despite differences, often quite wide, in orthography, lexicon, style, and pronunciation, they all reflect the same basic linguistic structure and a common if diverse literary tradition. The following approximate periods may be recognized for convenience: Early Latin, to the 2nd century B.C.; Classical Latin, divided into Republican, and Augustan ('Golden'), and Early Imperial ('Silver'), until the 2nd century A.D.; Late Latin (i.e., the Latin of Late Antiquity), 3rd to 6th century A.D.; Medieval Latin, from the 7th to the 13th century A.D., including two high points of Latin culture, the Carolingian period (9th century) and the 12th-century renaissance. The term *Neo-Latin* embraces Renaissance Latin (14th to 15th centuries), the scholarly and scientific Latin of the early modern period, and modern uses of the language.

Script

The Latin alphabet, the basis of current Western European alphabets, was derived from a form of Greek script introduced into Italy by colonists from Euboea. In its classical form, it contains all the letters that are still used in Modern English except for J, U, and W; V (the U shape developed later) was used for the vowel /u/ and the semivowel /w/; I was used for the vowel /i/ and the semivowel /j/. Y and Z, originally absent from Latin, were imported for use in Greek borrowings. The capital letters, used for inscriptions, have retained their classical form more or less unchanged. Originating in Roman cursive, a variety of handwritten styles evolved from late antiquity onward and ultimately gave rise to modern lowercase letters.

(ed.) *Dyslexia in different languages: cross-linguistic comparisons*. London: Whurr Publishers.
Yoshizaki K (2001). 'Effects of visual familiarity for words on interhemispheric cooperation for lexical processing.' *Cognitive Brain Research 12,* 409–414.
Young A W & Ellis A W (1985). 'Different methods of lexical access for words presented in the left and right visual hemifields.' *Brain and Language 24,* 326–358.
Zhang S & Perfetti C A (1993). 'The tongue-twister effect in reading Chinese.' *Journal of Experimental Psychology: Learning, Memory, and Cognition 19,* 1082–1093.
Zhu X (1987). 'Analysis of cueing function of phonetic components in modern Chinese.' In Yuan X (ed.) *Proceedings of the symposium on the Chinese language and characters*. Beijing: Guang Ming Daily Press.

Latin

J Powell, Royal Holloway, University of London, Egham, Surrey, UK

History and Affiliations

Latin is an Indo-European language of western central Italy, recorded from about 500 B.C. in the immediate area of Rome, related closely to the poorly attested Faliscan dialect and more distantly to Osco-Umbrian and Venetic. Lexical influences from neighboring languages (Italic dialects, the unrelated Etruscan language, and above all Greek) are visible in Latin from the earliest period. With the rise of Rome, Latin spread throughout Italy and into the provinces of the Roman Empire. It became the spoken language in western continental Europe, displacing the indigenous languages; its spoken varieties developed into the modern Romance languages. A more or less standard version of written Latin remained in use in administration, law, education, and the Church, and in due course spread throughout Europe as Medieval Latin. The status of Latin as a common medium of learned communication began to wane only with the rise of English as a world language in modern times.

Varieties

The dialect of Rome became the dominant form of Latin from an early period, with variations from it being regarded as rustic. Written Latin was standardized by the first century B.C. and remained broadly uniform thereafter, concealing changes and diversification in the spoken idiom; nevertheless a range of literary, technical, or colloquial registers can be distinguished in the surviving texts, and the occurrence of nonstandard written forms can provide evidence for spoken developments. The term *Vulgar Latin* is used to refer to any variety that departs significantly from classical norms. There is no good ground for regarding later written forms – Christian, Medieval, Renaissance, or Modern Latin – as essentially different varieties of the language: despite differences, often quite wide, in orthography, lexicon, style, and pronunciation, they all reflect the same basic linguistic structure and a common if diverse literary tradition. The following approximate periods may be recognized for convenience: Early Latin, to the 2nd century B.C.; Classical Latin, divided into Republican, and Augustan ('Golden'), and Early Imperial ('Silver'), until the 2nd century A.D.; Late Latin (i.e., the Latin of Late Antiquity), 3rd to 6th century A.D.; Medieval Latin, from the 7th to the 13th century A.D., including two high points of Latin culture, the Carolingian period (9th century) and the 12th-century renaissance. The term *Neo-Latin* embraces Renaissance Latin (14th to 15th centuries), the scholarly and scientific Latin of the early modern period, and modern uses of the language.

Script

The Latin alphabet, the basis of current Western European alphabets, was derived from a form of Greek script introduced into Italy by colonists from Euboea. In its classical form, it contains all the letters that are still used in Modern English except for J, U, and W; V (the U shape developed later) was used for the vowel /u/ and the semivowel /w/; I was used for the vowel /i/ and the semivowel /j/. Y and Z, originally absent from Latin, were imported for use in Greek borrowings. The capital letters, used for inscriptions, have retained their classical form more or less unchanged. Originating in Roman cursive, a variety of handwritten styles evolved from late antiquity onward and ultimately gave rise to modern lowercase letters.

Phonology

Sound/Spelling Relationship

The Classical Latin alphabet represents the sounds of the language quite well, except for its failure to distinguish the semivowels (as we now distinguish, e.g., *voluit* 'he wanted' from *volvit* 'he rolls'). Phonemic doubling of consonants was usually but not invariably indicated. In the classical period, vowel length was shown sometimes by doubling the vowel, more often by means of the *apex*, a diacritic resembling the acute accent. In Latin as conventionally printed, vowel length is not regularly indicated; so, for example, *canis* can represent either *cănĭs* 'dog' or *cānīs* 'grey hair (dative-ablative)'.

A number of sounds initially caused spelling difficulties: (a) Classical Latin had a series of nasal vowels, represented (consistently with etymology) by vowel + *m* finally and vowel + *n* medially; (b) a vowel sound midway between /i/ and /u/ was found in words such as *optumus/-imus* 'best' and other superlatives, and *lubens/libens* 'willing'; (c) there was no separate symbol for the velar nasal (see note *c* under the following consonant table).

Consonants

The consonantal inventory of Latin may be set out as follows:

	Labio-velar	*Velar*	*Dental*	*Labial*
Voiceless stop	/k^w/ <qu>	/k/[a] <c>/<k>	/t/	/p/
Voiced stop	/g^w/ <gu>[b]	/g/	/d/	/b/
Nasal		/ŋ/ <n>/<g>[c]	/n/	/m/
Fricative	glottal: /h/	alveolar: /s/		labiodental: /f/
Liquid		rolled/flapped: /r/ lateral: /l/[d]		
Semivowel		palatal: /j/ <i>		labial: /w/ <u>/<v>[e]

[a]The original sound was /k/, but the Romance and Medieval palatal realization before front vowels may already have started to develop in antiquity.
[b]Occurs only medially after a nasal or liquid as in *languidus* 'languid', *urgueo* 'press'.
[c]Occurs only medially, (a) as the outcome of /n/ before a velar, or (b) as the outcome of /c/ or /g/ before a nasal; written in the first case as <n>, as in *tango* 'touch', and in the second as <g> as in *dignus* 'worthy' (< **dec-nos*, cf. *dec-ens* 'decent'). The combination /ŋn/ <gn> later underwent a further sound change to palatal /ñ/, found, for example, in Italian.
[d]Represented by both 'clear' and 'dark' (velarized) allophones, as shown by the effect on neighboring vowels.
[e]The sound /w/ later changed to /β/ or /v/; spelling variations indicate that this took place (under Greek influence?) in Italy and the East possibly as early as the first century A.D., and in the western provinces at a later period.

Vowels

Classical Latin had a symmetrical five-way vowel system with phonemic distinction between long, short, and nasalized vowels: e.g., *puellă* 'girl' (nominative): *puellā* /-a:/ (ablative): *puellam* /-ã/ (accusative). Early Latin had a range of inherited diphthongs, which were considerably simplified by the Classical period; further leveling took place in varieties of the spoken language (*au* > /o:/, *ae* > open /ę:/, *oe* > /e:/), giving rise to confusion in spelling. In Romance, the classical vowel system was remodeled: nasalization and phonemic vowel length disappeared, while changes in vowel quality led to merger of short /i/ with long /e:/, and of short /u/ with long /o:/.

Accent and Prosody

Classical and post-Classical Latin has a largely predictable stress accent: the accent falls on a long penultimate, but on the antepenultimate if the penultimate is short. (A short syllable is an open syllable containing a short vowel; all other syllables, except for a few doubtful cases, are long.) This had apparently replaced an earlier initial stress accent, which is postulated to account for the regular vowel weakening that took place in early Latin in non-initial syllables, e.g., *pro* 'before' + *habeo* 'hold' > *prohibeo* 'forbid'. The usually clear distinctions of syllabic quantity made it possible for quantitative meters borrowed from Greek to replace the native 'Saturnian' verse form. Stress accent plays no formal part in Classical Latin verse, although there is an interaction between quantitative rhythm and word accent, which both creates some metrical restrictions and allows for considerable poetic subtlety. Stressed rhythms were reintroduced in Late and Medieval Latin.

Morphology

Inflection

Latin is an inflected language retaining a number of features of the parent Indo-European, but with considerable innovation, especially in the verbal paradigm.

Nouns are inflected for number and case. There are two numbers, singular and plural, and five regular cases: nominative, accusative, genitive, dative, and ablative. Nouns are traditionally grouped in five declensions: the first and second declensions continue the IE O/A stems, the third and fourth continue the IE stems in consonants and semivowels, and the fifth is a Latin innovation of mixed origin. A separate vocative is found only in the second declension and in borrowed Greek names. A locative case exists in place

names and in a few common nouns, such as *domi* 'at home' from *domus* 'home'. Nouns are assigned to one of three gender categories – masculine, feminine, and neuter – on the basis partly of sense and partly of form, so that on the whole there are fewer unpredictable or illogical assignments of gender than in many other languages.

Adjectives, both as attributes and as complements, agree with nouns in gender, number, and case; they broadly follow the same patterns of declension as nouns. There are no articles, but there is a rich pronominal system including four demonstratives (*is* 'the one mentioned', *hic* 'this', *iste* 'that of yours', *ille* 'that over there') and a distinction between specified and nonspecified indefinite pronouns (*quidam* 'someone in particular', *aliquis* 'someone or other').

Verbs are inflected according to the person and number of the subject. Subject pronouns are regularly dropped (*amamus* 'we love'). There are two sets of personal endings, indicating voice: active *-o/-m -s -t -mus -tis -nt*, passive *-r -ris/-re -tur -mur -mini -ntur*; and one other set of endings peculiar to the perfect active. Verbs are classified in four conjugations according to stem vowel. The principal parts of a Latin verb, from which all other forms can be derived, are the first person singular present indicative active (normally listed in the dictionary, e.g., *dico* 'I say'), the present infinitive active (e.g., *dicere* 'to say'), the first person singular perfect active (e.g., *dixi* 'I said'), and the past participle passive (e.g., *dictum* 'said'). The stems of the present and perfect are combined transparently with tense markers to form two further indicative tenses and two subjunctive from each stem, e.g., *proced-eba-m* 'I was proceeding' (imperfect indicative), *praedix-era-tis* 'you had predicted' (pluperfect indicative), *mane-re-mus* 'we would be staying' (imperfect subjunctive), *decipi-a-mini* 'you would be deceived' (present subjunctive passive). There are two imperatives, one used for immediate commands (*exi!* 'get out!'), the other for instructions (*bene misceto* 'mix well'). The perfect tenses of the passive are formed periphrastically: the past participle is combined with an appropriate (present-stem) form of *esse* 'be', e.g., *occisa est* 'she has been killed'.

A restricted class of verbs (generally known as 'deponents': fossilized remnants of the old middle voice) take passive morphology although their meaning is active and can be transitive: e.g., *philosophor* 'I philosophize'; *vereor* (transitive) 'I fear'. Intransitive active verbs (including a class of verbs that take a dative object) have no passive, except third-person singular forms used impersonally, e.g., *venio* 'I arrive'; *venitur* 'people arrive' (cf. the impersonal passive in the Celtic languages). Genuinely irregular verbs are few (*sum* 'be', *possum* 'can', *volo* 'want', and some others). The greatest difficulty in Latin verbal morphology is the often unpredictable formation of the perfect stem.

The verb has two infinitives, present (*dicere* 'to say') and perfect (*dixisse* 'to have said'); other verbal nouns (gerunds, supines) supply case forms for the infinitive. There are three participles, present active (*dicens* 'saying'), past passive (*dictum* 'said'), and future active (*dicturus* 'about to say'), and a verbal adjective (the gerundive), which is future passive in meaning and often denotes obligation (*faciendum est* 'it is to be done').

Derivational Morphology

Latin has a productive system of derivational morphology, much of which is familiar through derivatives in modern languages, e.g., prepositional prefixes such as *ex-* 'out', *circum-* 'round'; other prefixes such as *in-* 'not', *re-* 'back', nominal suffixes such as *-tio(n)-* (process) and *-tor* and *-trix* (male and female agents), and adjectival suffixes such as *-bilis* (*inextricabilis* 'inextricable'), *-anus* (*Christianus* 'Christian'), *-arius* (*piscarius* 'to do with fish'). Adjectives form adverbs (in *-e*, *-o*, or *-ter*), comparatives in *-ior*, and superlatives in *-issimus*. At first, Latin was resistant to some kinds of derivation; for example, compound nouns or adjectives were relatively unusual, occurring mostly in high style (*solivagus* 'wandering alone') or as comic colloquialisms (*caldicerebrius* 'hot-headed'). From the classical period onward, the need to express new ideas led to a large expansion of the classical lexicon both by derivation and by borrowing.

Syntax

Word Order

Because Latin encodes basic grammatical relations by means of inflection, there is considerable freedom in word order. Either a topic or an emphasized predicate may be placed first in a sentence, and variations of logical emphasis are also common at phrase level. Classical Latin preserves the Indo-European phenomenon known as Wackernagel's Law, i.e., the first stressed position in a sentence or clause is followed by an unstressed position which often contains, in order of precedence, (a) sentence connectives, (b) weak pronouns, (c) unstressed verbs. The phenomenon of hyperbaton, or discontinuity in the noun phrase, is a particular feature of Latin as it is of Classical Greek: its effect is often to

throw particular focus on to the first element of the noun phrase, e.g.:

bonos	*habemus*	*consules*
good-MASC.ACC.PL. + Focus	we-have	consuls-ACC.PL.

'we have *good* consuls'

Because of inflection, the absence of articles, and the flexibility of word order, Latin is capable of considerable compression, hence the lapidary style favored in inscriptions, proverbs, and epigrams.

Complex Sentence Structure

Subordination is much favored in varieties of literary Latin; Classical Roman writers influenced by the Greek periodic style made full use of the available syntactical resources, e.g., the inflected relative pronoun enabling any nominal constituent to be relativized, and the free use of the accusative-infinitive construction to create continuous indirect speech (*oratio obliqua*), as well as paratactic devices such as apposition, parallelism, ellipsis, and chiasmus. Means were found to compensate for the lack of certain features (e.g., the past participle active) which had made for flexibility in Greek, and the rich tense system is fully exploited in classical narrative prose. Late and Medieval Latin usage often blurs the distinctions of tense and mood found in Classical Latin; the conjunction *quod* 'that' gains ground at the expense of the accusative-infinitive, the active present participle becomes more frequent, and the periphrastic constructions characteristic of the Romance verbal conjugation begin to appear.

See also: Greek, Ancient; Indo–European Languages; Latin Lexicography; Romance Languages.

Bibliography

Adams J N (2003). *Bilingualism and the Latin language.* Cambridge: Cambridge University Press.

Allen W S (1978). *Vox Latina. The pronunciation of Classical Latin* (2nd edn.). Cambridge: Cambridge University Press.

Baldi P (1999). *The foundations of Latin.* (Trends in Linguistics: Studies and Monographs 117.) Berlin and New York: Mouton de Gruyter.

Courtney E (1999). *Archaic Latin prose.* (American Philological Association: American Classical Studies, 42.) Atlanta: Scholars Press.

Cupaiuolo F (1993). *Bibliografia della lingua latina 1949–1991.* Naples: Loffredo.

De Meo C (1986). *Lingue tecniche del latino* (2nd edn.). Bologna: Pàtron.

Ernout A (1953). *Morphologie historique du latin* (3rd edn.). Paris: Klincksieck.

Ernout A (1954). *Aspects du vocabulaire latin.* Paris: Klincksieck.

Glare P G W (ed.) (1982). *Oxford Latin dictionary.* Oxford: Oxford University Press.

Hammond M (1976). *Latin: a historical and linguistic handbook*. Cambridge, MA, and London: Harvard University Press.

Herman J (1997). *Vulgar Latin.* Wright R (English trans.). University Park, PA: The Pennsylvania State University Press.

Hofmann J B (1951). *Lateinische Umgangssprache* (3rd edn.). Heidelberg: C. Winter.

Kennedy B H (1962). *The Revised Latin primer.* Ed. and rev. Mountford, J. London: Longmans.

Keppie L (2001, 1991). *Understanding Roman inscriptions.* London: Routledge.

Löfstedt E (1959). *Late Latin.* Oslo: Aschehoug; Cambridge, MA: Harvard University Press.

Marouzeau J (1953). *L'Ordre des mots en latin* (new edn.) *(Collection d'études latines, série scientifique)* (3 vols). Paris: Les Belles Lettres.

Marouzeau J (1962). *Traité de stylistique latine* (4th edn.). Paris: Les Belles Lettres.

Meillet A (1977). *Esquisse d'une histoire de la langue latine.* (3rd edn., rev. Perrot J). Paris: Klincksieck.

Meiser G (1998). *Historische Laut- und Formenlehre der lateinischen Sprache.* Darmstadt: Wissenschaftliche Buchgesellschaft.

Mohrmann C (1961–1965). *Études sur le latin des chrétiens* (3 vols). Rome: Edizioni di Storia e Letteratura.

Niedermann M (1953). *Historische Lautlehre des Lateinischen* (3rd edn.). Heidelberg: C. Winter.

Norberg D (1968). *Manuel pratique de Latin Médiéval.* Paris: A. & J. Picard.

Palmer L R (1968). *The Latin language.* London: Faber and Faber.

Pinkster H (1990). *Latin syntax and semantics*. London: Routledge.

Poccetti P, Poli D & Santini C (1999). *Una storia della lingua latina. Formazione, usi, comunicazione.* Rome: Carocci.

Raven D S (1965). *Latin metre: An introduction.* London: Faber and Faber.

Rosén H (1999). *Latine loqui. Trends and directions in the crystallisation of Classical Latin.* Munich: Wilhelm Fink Verlag.

Sandys J E (ed.) (1921). *A companion to Latin studies* (3rd edn.). Cambridge: Cambridge University Press.

Stolz F, Debrunner A & Schmid W P (1993). *Storia della lingua latina.* Italian tr. Benedikter C, intr. and notes Traina A (4th edn., rev. Vineis E). Bologna: Pàtron.

Woodcock E (1959). *A new Latin syntax*. London: Methuen.

Latin Lexicography

E Stathi, Berlin-Brandenburg Academy of Sciences, Berlin, Germany

Ancient and Medieval Lexicography

Roman contact with the Greek world and familiarity with Greek scholarship led to the creation of bilingual Greek–Latin and Latin–Greek word lists and dictionaries for practical and educational uses in the classical period. Monolingual Latin lexicography can be seen in the few works that have survived, e.g., that of Nonius, a grammarian of the 3rd and 4th centuries, and Isidor, bishop of Seville (7th century), in the tenth book of his encyclopedia *Etymologiae*. Verrius Flaccus' work, *De verborum significatu,* is preserved in an abridged version by Festus and was further shortened by Paulus Diaconus. Furthermore, glossaries, mostly anonymous, served educational purposes, explaining difficult words in literature. One of the most important collections is the *Liber glossarium* (8th century). In early medieval times this served as input for two important works, Papias' *Elementarium* (11th century) and Osbern of Gloucester's *Derivationes* (12th century). These two were the precursors of late medieval lexicography as manifested in the *Liber derivationum* by Hugutio and the *Catholicon* by Johannes Januensis.

Modern Latin Lexicography

During the Renaissance much lexicography was devoted to explaining Latin terms in the various vernacular languages of Europe. Two major monolingual Latin lexicographic works appeared. The *Dictionarium* (1502) by Ambrosius Calepinus, the Augustinian monk of Bergamo, was innovative in that the etymology was seen as secondary to the explanation of meanings. Even more importantly, examples of the uses of words are given. Calepinus made a second edition, and after his death his dictionary was revised and reprinted many times. In Paris, Robertus Stephanus (Robert Estianne)'s plan to produce a new edition of Calepinus resulted in a totally new dictionary in 1531, the *Dictionarium seu Latinae linguae thesaurus*, based on Calepinus and on citations from literature. The French glosses were removed in the revised edition (1536) to make room for a huge amount of additional Latin material. This led to the first compendious Latin monolingual dictionary, which remained the authority until the late 18th century.

A new era of Latin lexicography was inaugurated by Egidio Forcellini's *Totius latinitatis lexicon* (1771), which offers a huge collection of materials. An innovative contribution was the internal structure of the entries: Forcellini distinguished strictly between different meanings by exhibiting the relation between the core meaning and the derived meanings of a word (e.g., figurative usage). This dictionary, which also takes account of inscriptions, was abridged and revised by many scholars and was the basis for most 19th-century dictionaries. The abridgment by Wilhelm Freund (1834–1835) served as foundation for the dictionary by E. A. Andrews (1850), which in turn was revised and reedited by Charlton T. Lewis and Charles L. Short (1879).

The project *Thesaurus linguae Latinae (TLL)*, now hosted by the Bavarian Academy of Sciences in Munich, is compiling a scholarly dictionary that covers all of existing Latin literature from the beginnings until 600 A.D. The first fascicle was published in 1900; by 2004, the word *propositio* had been reached. P. G. W. Glare's *Oxford Latin dictionary* (1968–1975) is the most extensive and complete dictionary available today. It includes literary, epigraphic, and other sources. Still valuable is K.-E. Georges' *Ausführliches lateinisch-deutsches Handwörterbuch* (revised by H. Georges, 2 vol., 1913–1918, reprinted in 1988). The Perseus Project at Tufts University is an online dictionary with English metalanguage based on Lewis and Short.

Medieval Latin

A largely neglected area of Latin scholarship, medieval Latin was first treated by Du Cange's *Glossarium mediae et infimae latinitatis* (1678), which covers the period from 800 to 1200 and is still used as a reference for later Latin. The project *Das mittellateinische Wörterbuch*, hosted at the Bavarian Academy of Sciences, covers the whole vocabulary of medieval Latin (even words that occurred already in antiquity). It covers the period from 600 A.D. (where the *TLL* stops) until 1280. The definitions are given in German. By 2004 the work had reached the word *dominum*.

Etymological dictionaries

The two most authoritative etymological dictionaries of Latin are the *Dictionnaire étymologique de la langue latine* (1932, fourth edition 1959) by Alfred Ernout and Antoine Meillet and the *Lateinisches*

etymologisches Wörterbuch by Alois Walde (1906), revised by Johann B. Hofmann (third edition 1954–1956).

See also: Computational Lexicons and Dictionaries; Corpora; Du Cange, Charles du Fresne, Sieur (1610–1688); Encyclopédie; Etymology; Greek Lexicography, Classical; Latin; Lexicography: Overview; Stephanus, Robertus (ca. 1503–1559); Thesauruses.

Bibliography

Krömer D (1990). 'Lateinische lexikographie.' In Hausmann F J *et al.* (eds.) *Dictionaries: an international encyclopedia of lexicography.* Berlin/New York: de Gruyter. 1713–1722.

Latvia: Language Situation

I Druviete, University of Latvia, Riga, Latvia

Latvia [Latvija] is one of the three Baltic states. Its territory covers about 64 000 km^2. It has approximately 2 309 000 inhabitants. The Republic of Latvia was founded in 1918; its independence was renewed in 1991. In March 2004 it joined NATO, and as of May 1, 2004, it is a full member of the European Union.

Around 3000 B.C., the Finno-Ugric Livs settled there. During the first half of the next millennia, the ancient Baltic tribes arrived. The formation of the common Latvian language began during the 10th to 12th centuries A.D., when German crusaders conquered the territory. In the 16th century, the Livonian nation collapsed and today's Latvian territory fell into the hands of the Polish-Lithuanian Commonwealth until the Polish-Swedish War. After the Great Northern War (1700–1721) today's Latvian territory became part of the Russian Empire. In 1918, Latvia declared its independence. On June 17, 1940, Latvia was occupied by the Soviet Army. Latvia was again recognized as an independent state on August 21, 1991.

The ethnic composition in Latvia is as follows: 58.7% Latvians, 28.8% Russians, 3.9% Belorussians, 2.6% Ukrainians. 2.5% Poles, 1.4% Lithuanians, 0.4% Jews, 0.3% Roma, 0.2% Germans, 1.2% others (2004 Civil Data Register).

The Latvian language belongs to the Baltic group of the Indo-European family of languages. It is spoken by approximately 1.55 million people in Latvia and abroad as the first language and by approximately 500 000 people in Latvia as the second language. Typologically, Latvian is an inflectional language. The first written texts in Latvian appeared in the 16th century.

The Latvian language is the official language in the Republic of Latvia. The Law on the State Language (1999) is in force now. Ethnic minorities have the right to preserve and develop their language and their identity. The state finances primary education in eight languages.

The Liv language has a particular status of the only autochthonous minority language. Though about 170 people consider themselves Livs today, there are only five native speakers left.

The Russian language is the native language for 665 000 Russians, due to Russification processes for 72.8% Belorussians, 67.8% Ukrainians, and 57.7% Poles in Latvia. The first Russian settlers came to Latvia during 16th and 17th centuries. In 1935, 8.8% of the population in Latvia were Russian; during the Soviet period the number of Russians increased more than four times. Russian now is the language of tuition for 30% of all students. Apart from the publications brought from Russia, 70 local newspapers are on offer. Several TV channels and radio stations broadcast in Russian.

The Polish language was widespread in the eastern part of Latvia, yet today its role has diminished. Of the 60 000 Latvian Poles, only 11 500 recognize it as their native language. There are six Polish schools in Latvia. The Belorussian language is known by 36 000 Latvian residents, yet it is rarely used at home. There is one Belorussian school in Riga with about 100 students. Ukrainians mainly entered Latvia after World War II. The Ukrainian secondary school has about 300 students. Small Estonian, Lithuanian, and German communities have active cultural lives. Romany in Latvia has a particular dialect. A language primer and various books have been published in it.

See also: Estonia: Language Situation; Lithuania: Language Situation.

etymologisches Wörterbuch by Alois Walde (1906), revised by Johann B. Hofmann (third edition 1954–1956).

See also: Computational Lexicons and Dictionaries; Corpora; Du Cange, Charles du Fresne, Sieur (1610–1688); Encyclopédie; Etymology; Greek Lexicography, Classical; Latin; Lexicography: Overview; Stephanus, Robertus (ca. 1503–1559); Thesauruses.

Bibliography

Krömer D (1990). 'Lateinische lexikographie.' In Hausmann F J *et al.* (eds.) *Dictionaries: an international encyclopedia of lexicography.* Berlin/New York: de Gruyter. 1713–1722.

Latvia: Language Situation

I Druviete, University of Latvia, Riga, Latvia

Latvia [Latvija] is one of the three Baltic states. Its territory covers about 64 000 km^2. It has approximately 2 309 000 inhabitants. The Republic of Latvia was founded in 1918; its independence was renewed in 1991. In March 2004 it joined NATO, and as of May 1, 2004, it is a full member of the European Union.

Around 3000 B.C., the Finno-Ugric Livs settled there. During the first half of the next millennia, the ancient Baltic tribes arrived. The formation of the common Latvian language began during the 10th to 12th centuries A.D., when German crusaders conquered the territory. In the 16th century, the Livonian nation collapsed and today's Latvian territory fell into the hands of the Polish-Lithuanian Commonwealth until the Polish-Swedish War. After the Great Northern War (1700–1721) today's Latvian territory became part of the Russian Empire. In 1918, Latvia declared its independence. On June 17, 1940, Latvia was occupied by the Soviet Army. Latvia was again recognized as an independent state on August 21, 1991.

The ethnic composition in Latvia is as follows: 58.7% Latvians, 28.8% Russians, 3.9% Belorussians, 2.6% Ukrainians. 2.5% Poles, 1.4% Lithuanians, 0.4% Jews, 0.3% Roma, 0.2% Germans, 1.2% others (2004 Civil Data Register).

The Latvian language belongs to the Baltic group of the Indo-European family of languages. It is spoken by approximately 1.55 million people in Latvia and abroad as the first language and by approximately 500 000 people in Latvia as the second language. Typologically, Latvian is an inflectional language. The first written texts in Latvian appeared in the 16th century.

The Latvian language is the official language in the Republic of Latvia. The Law on the State Language (1999) is in force now. Ethnic minorities have the right to preserve and develop their language and their identity. The state finances primary education in eight languages.

The Liv language has a particular status of the only autochthonous minority language. Though about 170 people consider themselves Livs today, there are only five native speakers left.

The Russian language is the native language for 665 000 Russians, due to Russification processes for 72.8% Belorussians, 67.8% Ukrainians, and 57.7% Poles in Latvia. The first Russian settlers came to Latvia during 16th and 17th centuries. In 1935, 8.8% of the population in Latvia were Russian; during the Soviet period the number of Russians increased more than four times. Russian now is the language of tuition for 30% of all students. Apart from the publications brought from Russia, 70 local newspapers are on offer. Several TV channels and radio stations broadcast in Russian.

The Polish language was widespread in the eastern part of Latvia, yet today its role has diminished. Of the 60 000 Latvian Poles, only 11 500 recognize it as their native language. There are six Polish schools in Latvia. The Belorussian language is known by 36 000 Latvian residents, yet it is rarely used at home. There is one Belorussian school in Riga with about 100 students. Ukrainians mainly entered Latvia after World War II. The Ukrainian secondary school has about 300 students. Small Estonian, Lithuanian, and German communities have active cultural lives. Romany in Latvia has a particular dialect. A language primer and various books have been published in it.

See also: Estonia: Language Situation; Lithuania: Language Situation.

Bibliography

Apine I & Volkovs V (1998). *Slāvi Latvijā (Etniskās vēstures apcerējums)*. Rīga: Mācību apgāds.

Blinkena A (1994/1995). 'The Latvian language: some problems of its development and existence.' *Actes de langage française et de linguistique: ALFA 1994/95 7/8*, 463–469.

Dreifelds J (1996). *Latvia in transition*. Cambridge: Cambridge University Press.

Dribins L (1994). 'A nation state and its ethnic policies.' *Humanities and Social Sciences: Latvia 2(3),* 70–93.

Dribins L (ed.) (1996). *The national and ethnic groups in Latvia*. Rīga: Preses Nams.

Druviete I (1997). 'Linguistic human rights in the Baltic States.' *International Journal of the Sociology of Language 127,* 161–185.

Druviete I (1997). 'Latvia.' In Goebl H, Nelde P H & Stary Z (eds.) *Contact linguistics: an international handbook of contemporary research, 2*. Berlin/New York: Walter de Gruyter. 1906–1912.

Druviete I (1998). 'Republic of Latvia.' In Bratt Paulston C & Peckham D (eds.) *Linguistic minorities in the Central and Eastern Europe*. Clevedon: Multilingual Matters. 160–183.

Druviete I (1998). 'La situation sociolinguistique de la langue lettone.' *Terminogramme. Juillet 1998. Les politiques linguistiques des pays baltes.* 105–149.

Druviete I (1999). 'Language policy in a changing society: problematic issues of implementation of universal linguistic human rights standards.' In Kontra M, Phillipson R & Skutnabb-Kangas T (eds.) *Approaching linguistic human rights*. Budapest: Central European University Press. 263–276.

Druviete I (2000). 'The Latvian language law debate: some aspects of linguistic human rights in education.' In Phillipson R (ed.) *Rights to language: equity, power, and education: celebrating the 60th birthday of Tove Skutnabb-Kangas*. Mahwah, NJ: Lawrence Erlbaum Associates. 83–86.

Druviete I (2000). 'La Charte de la langue française et les lois linguistiques dans les pays Baltes.' *Terminogramme. Revue d'amenagement linguistique. Hors serie*. 217–221.

Druviete I (2002). 'The future of the Latvian language in the enlarged European Union.' *Humanities and Social Sciences: Latvia 1(34),* 34–46.

Jubulis M A (2001). *Nationalism and democratic transition: the politics of citizenship and language in post-Soviet Latvia*. Lanham: University Press of America.

Karklins R (1994). *Ethnopolitics and transition to democracy: the collapse of the USSR and Latvia*. Washington, D.C.: The Woodrow Wilson Center Press.

Loeber A (1993). 'Language rights in independent Estonia, Latvia and Lithuania, 1918–1940.' In Vilfan S (ed.) *Ethnic groups and language rights, 3*. New York: New York University Press. 221–249.

Ozolins U (1994). 'Upwardly mobile languages: the politics of language in the Baltic States.' *Journal of Multilingual and Multicultural Development 15(2–3),* 161–170.

Ozolins U (1999). 'Between Russian and European hegemony: current language policy in the Baltic States.' *Current Issues in Language & Society 6(1),* 6–47.

Rannut M (1994). 'Beyond linguistic policy: the Soviet Union versus Estonia.' In Skutnabb-Kangas T, Phillipson R & Rannut M (eds.) *Linguistic human rights. Overcoming linguistic discrimination.* Berlin/New York: Mouton de Gruyter. 179–208.

Valdmanis J (1997). 'Processes determining the future of the Latvian language.' In *Proceedings of the European conference on language planning*. Barcelona: Departament de Cultura. 45–49.

Zvidriņš P (1993). 'Changes of ethnic structure in Latvia.' *Humanities and Social Sciences: Latvia 1(1),* 10–26.

Latvian

S Young, University of Maryland Baltimore County, Baltimore, MD, USA

Latvian (*latviešu valoda*), with some 1.5 million speakers in the Republic of Latvia, is one of two present-day Indo-European Baltic languages, the other being the structurally more conservative Lithuanian. The Latvian standard language is based on the central Latvian dialect (*vidus dialekts*), which is further divided (from west to east) into Couronian, Zemgalian, and Vidzeme varieties. The central dialect, together with Tamian in northwestern Couronia and along the northeastern coast of the Bay of Riga, is known as Low Latvian. High Latvian (Selonian and Latgalian, the latter with an independent literary tradition), is found in the eastern third of the country.

The traditional view of the origin of Latvian is that it represents a synthesis of the language of the early Latgalians (known from the 13th century on simply as *Letti,* that is, Letts or Latvians) and neighboring closely related Baltic dialects, now extinct, among them Zemgalian (along the Lielupe river), Couronian (in southwestern Latvia), and Selonian (along the middle Daugava). The influence of these Baltic substrata, together with that of Balto-Finnic and Slavic neighbors, accounts in large part for the marked

Bibliography

Apine I & Volkovs V (1998). *Slāvi Latvijā (Etniskās vēstures apcerējums).* Rīga: Mācību apgāds.

Blinkena A (1994/1995). 'The Latvian language: some problems of its development and existence.' *Actes de langage française et de linguistique: ALFA 1994/95 7/8*, 463–469.

Dreifelds J (1996). *Latvia in transition.* Cambridge: Cambridge University Press.

Dribins L (1994). 'A nation state and its ethnic policies.' *Humanities and Social Sciences: Latvia 2(3),* 70–93.

Dribins L (ed.) (1996). *The national and ethnic groups in Latvia.* Rīga: Preses Nams.

Druviete I (1997). 'Linguistic human rights in the Baltic States.' *International Journal of the Sociology of Language 127,* 161–185.

Druviete I (1997). 'Latvia.' In Goebl H, Nelde P H & Stary Z (eds.) *Contact linguistics: an international handbook of contemporary research, 2.* Berlin/New York: Walter de Gruyter. 1906–1912.

Druviete I (1998). 'Republic of Latvia.' In Bratt Paulston C & Peckham D (eds.) *Linguistic minorities in the Central and Eastern Europe.* Clevedon: Multilingual Matters. 160–183.

Druviete I (1998). 'La situation sociolinguistique de la langue lettone.' *Terminogramme. Juillet 1998. Les politiques linguistiques des pays baltes.* 105–149.

Druviete I (1999). 'Language policy in a changing society: problematic issues of implementation of universal linguistic human rights standards.' In Kontra M, Phillipson R & Skutnabb-Kangas T (eds.) *Approaching linguistic human rights.* Budapest: Central European University Press. 263–276.

Druviete I (2000). 'The Latvian language law debate: some aspects of linguistic human rights in education.' In Phillipson R (ed.) *Rights to language: equity, power, and education: celebrating the 60th birthday of Tove Skutnabb-Kangas.* Mahwah, NJ: Lawrence Erlbaum Associates. 83–86.

Druviete I (2000). 'La Charte de la langue française et les lois linguistiques dans les pays Baltes.' *Terminogramme. Revue d'amenagement linguistique. Hors serie.* 217–221.

Druviete I (2002). 'The future of the Latvian language in the enlarged European Union.' *Humanities and Social Sciences: Latvia 1(34),* 34–46.

Jubulis M A (2001). *Nationalism and democratic transition: the politics of citizenship and language in post-Soviet Latvia.* Lanham: University Press of America.

Karklins R (1994). *Ethnopolitics and transition to democracy: the collapse of the USSR and Latvia.* Washington, D.C.: The Woodrow Wilson Center Press.

Loeber A (1993). 'Language rights in independent Estonia, Latvia and Lithuania, 1918–1940.' In Vilfan S (ed.) *Ethnic groups and language rights, 3.* New York: New York University Press. 221–249.

Ozolins U (1994). 'Upwardly mobile languages: the politics of language in the Baltic States.' *Journal of Multilingual and Multicultural Development 15(2–3),* 161–170.

Ozolins U (1999). 'Between Russian and European hegemony: current language policy in the Baltic States.' *Current Issues in Language & Society 6(1),* 6–47.

Rannut M (1994). 'Beyond linguistic policy: the Soviet Union versus Estonia.' In Skutnabb-Kangas T, Phillipson R & Rannut M (eds.) *Linguistic human rights. Overcoming linguistic discrimination.* Berlin/New York: Mouton de Gruyter. 179–208.

Valdmanis J (1997). 'Processes determining the future of the Latvian language.' In *Proceedings of the European conference on language planning.* Barcelona: Departament de Cultura. 45–49.

Zvidriņš P (1993). 'Changes of ethnic structure in Latvia.' *Humanities and Social Sciences: Latvia 1(1),* 10–26.

Latvian

S Young, University of Maryland Baltimore County, Baltimore, MD, USA

Latvian (*latviešu valoda*), with some 1.5 million speakers in the Republic of Latvia, is one of two present-day Indo-European Baltic languages, the other being the structurally more conservative Lithuanian. The Latvian standard language is based on the central Latvian dialect (*vidus dialekts*), which is further divided (from west to east) into Couronian, Zemgalian, and Vidzeme varieties. The central dialect, together with Tamian in northwestern Couronia and along the northeastern coast of the Bay of Riga, is known as Low Latvian. High Latvian (Selonian and Latgalian, the latter with an independent literary tradition), is found in the eastern third of the country.

The traditional view of the origin of Latvian is that it represents a synthesis of the language of the early Latgalians (known from the 13th century on simply as *Letti*, that is, Letts or Latvians) and neighboring closely related Baltic dialects, now extinct, among them Zemgalian (along the Lielupe river), Couronian (in southwestern Latvia), and Selonian (along the middle Daugava). The influence of these Baltic substrata, together with that of Balto-Finnic and Slavic neighbors, accounts in large part for the marked

dialectal diversity of Latvian within a relatively small territory.

The beginnings of a Latvian literary tradition date to the Reformation. The first published book in Latvian is a 1585 translation of a Catholic catechism; this was soon followed by other religious texts, chiefly translations. The language of these early texts is to varying degrees influenced by the German speech of the authors, especially in syntax. Efforts to establish a national standard language were begun in the mid-19th century; the Latvian traditional folksongs (*dainas*) served as an important source of norms for the standard language. A milestone in the codification and description of modern Latvian was the appearance of Jan Endzelin's (Jānis Endzelīns) *Lettische grammatik* (1922) and the four-volume *Latviešu valodas vārdnīca* [*Dictionary of Latvian*] (1923–32), begun by Karl Mühlenbach (Kārlis Mīlenbahs) and completed by Jan Endzelin.

The Latvian vowel phonemes include /i, u, ɛ, æ, a/ (spelled *i, u, e, e, a*) and their long counterparts /iː, uː, ɛː, æː, aː/ (spelled *ī, ū, ē, ē, ā*). In addition, a long and short *o* are found in borrowed words. The vowels /ɛ/ and /æ/ (and /ɛː/ and /æː/) are represented in writing by a single *e* (*ē*), reflecting the fact that they are derived from a single source, Baltic *e* (**ē*). Originally conditioned by the nature of the following vowel, they now function as separate phonemes. Functioning as long vowels are the diphthongs [i͜ɐ, u͜ɐ] (written *ie, o*), which result from Common Baltic **ei* and **ō*: *dìevs* 'God' (**dei̯u-*), *nò* 'from' (**nō*), and from tautosyllabic sequences *en* and *an*: *pìeci* [pì͜ɐt͜si] 'five,' *ròka* [rù͜ɐka] 'hand' (Li. *penkì, rankà*). In final syllables Baltic long vowels are shortened and short *e* and *a* are lost: La. *vìlks* 'wolf': Li. *vil̃kas*.

Among the consonants of Latvian, the obstruents reflect a voiced:voiceless opposition, with regressive voicing assimilation (*atbilde* ['adbilde] 'answer,' *labs* [laps] 'good'); unlike Lithuanian, word-final consonants in final position are not devoiced: *kad* [kad] 'when.' Latvian lacks the feature of palatalization found in Lithuanian, but distinguishes a series of palatal consonants: *ķ* [c], *ģ* [ɟ], *ņ* [ɲ], *ļ* [ʎ]; *ķ* and *ģ* are found mainly in borrowed words. Characteristic of the Latvian consonant system is the development of t͜s, d͜z from Baltic *k, g* before a front vowel: *cits* 'other' (Li. *kìtas*), *dzìmt* 'to be born' (Li. *gim̃ti*). The Baltic reflexes of the IE palatovelars have merged in Latvian with *s* and *z*: *sìmts* 'hundred' (Li. *šim̃tas*) < IE **ḱm̥tó-*, *zeme* 'earth' (Li. *žemė*) < IE **dʰǵʰem-*. Secondary palatals ʃ, ʒ (orthographically *š, ž*) have arisen in turn from the sequences **si̯*, **ti̯* (>ʃ) and **zi̯*, **di̯* (>ʒ) before a back vowel: *šũt* 'to sew' (<**si̯ūt-*).

With a few exceptions, Latvian has initial stress. The standard language distinguishes three phonemic tones (marked only in linguistic texts) on long vowels and diphthongs: level [˜] and broken [ˆ], both representing Baltic acute tone; and falling [`], representing Baltic circumflex: *mĩt* 'to tread,' *mît* 'to change (arch.),' *mìt* 'resides.' Only the central Latvian Vidzeme dialect around Valmiera still distinguishes all three such tones; remaining dialects oppose only two (outside of High Latvian, broken and falling tone typically merge as either falling or broken tone).

The Latvian noun distinguishes masculine and feminine gender, each with three declensional patterns. Within declensional paradigms, five cases are distinguished morphologically: nominative, genitive, dative, accusative, and locative, in both the singular and plural. Adjectives, which occur in both definite and indefinite forms, agree with the noun in number, gender, and case.

The Latvian verb distinguishes present, past, and future tenses, and has a system of relative tenses formed with *būt* 'to be' and the past active participle. The verb has three conjugational types (only two of which are productive), each with a number of subtypes. The 1 sg. present and past forms coalesce in many instances where the stems are not distinguished by root vowel gradation. Like Lithuanian, singular and plural are not distinguished morphologically in the third person.

Peculiar to Latvian, from a Baltic perspective, is the syntactic construction *man ir* 'to-me is' for 'I have' (Li. *aš turiù*) and the use of the debitive (expressing obligation), formed with the particle *jā-* and the 3rd person present, with a dative subject: *man jābrauc* ("to-me must-go") 'I must go.' Like Lithuanian, Latvian uses a preposed adnominal genitive: *latviešu valoda* ("of-Latvians language" = 'Latvian language').

The lexicon has been rather strongly influenced by neighboring Baltic Finnic languages, chiefly Livonian: interrogative particle *vài* (Liv. *või*), *mãja* 'house' (Est. *maja*); and also by German: Middle Low German during the Hanseatic period and New High German thereafter (*brīvs* 'free': MLG *vrī*, *ùn* 'and': MLG *un*, *bril̃les* 'eyeglasses' NHG *die Brille*). Nevertheless, Latvian preserves a number of Indo-European archaisms not found elsewhere in Baltic: *asins* 'blood,' *agrs* 'early.'

See also: Balto-Slavic Languages; Latvia: Language Situation; Latvian Lexicography; Lithuanian.

Bibliography

Dini P (2000). *Baltų kalbos: lyginamoji istorija.* Vilnius: Mokslo ir enciklopedijų leidybos institutas.

Endzelin J (1922). *Lettische Grammatik.* Riga: A. Gulbis.

Fennell T G & Giesen H (1980). *A grammar of modern Latvian* (3 vols). The Hague: Mouton.

Forssman B (2001). *Lettische grammatik*. Dettelbach: J. H. Röll.
Gāters A (1977). *Die lettische Sprache und Ihre Dialekte*. The Hague: Mouton.
Kabelka J (1982). *Baltų filologijos įvadas*. Vilnius: Mokslas.
Mathiassen T (1997). *A short grammar of Latvian*. Columbus, OH: Slavica.
Mīlenbahs K & Endzelīns J (1923–1932). *Latviešu valodas vārdnīca* (4 vols). Riga: Kultūras fonda izdevums.
Nau N (1998). *Latvian*. Munich and Newcastle: Lincom Europa.
Rūķe-Draviņa V (1977). *The standardization process in Latvian: 16th century to the present*. Stockholm: Almqvist & Wiksell International.

Latvian Lexicography

A Veisbergs, University of Latvia, Riga, Latvia

Latvian lexicography has been and remains primarily bilingual. This has been historically determined by the dominance of one or another country in the region – so that dictionaries accommodated to the situation accordingly. It is usual to date Latvian lexicography from 1638, when G. Mancelius's *Lettus* was published, containing about 6000 German-Latvian entries. There were also two multilingual dictionaries published in the same century: a Polish-Latin-Latvian work by Elger (1683) and a German-Latin-Polish-Latvian one by Dressel (1688). The following two centuries saw the same pattern – although the dictionaries increased in size and quality, they were made by non-Latvians. The situation changed in the middle of the 19th century, when after the national 'awakening' Latvian linguists took hold of lexicography. The new dictionaries were purged of German and Russian loanwords, and they came closer to spoken language yet also became more prescriptive, the Waldemahr (1872), with 50 000 entries. The dictionary scene up to 1900 was thoroughly researched by Zemzare (1961).

The iconic Latvian dictionary of the 20th century is the four-volume work by Muehlenbach and Endzelins, published in 1923, with supplements appearing in 1932 and 1946. It contains a wealth of Latvian word-stock, though purged of later loanwords. The dictionary contains roughly 120 000 entries in the new orthography with explanations (in German), etymologies, stress patterns, citations, and examples and, despite the passage of time and several shortcomings, it is still viewed as the 'proper' source of Latvian lexis.

The beginning of the 20th century saw an expansion of bilingual dictionaries in the new contact languages – English and French – as well as a strong emphasis on dictionaries of foreign words. The Soviet period produced the first extensive monolingual explanatory dictionary of literary Latvian (*Latviešu literārās valodas vārdnīca*) in eight volumes, which, however, suffers from a lack of colloquial language and is affected by ideological constraints. A one-volume desk dictionary (25 000 entries) of standard Latvian, *Latviešu valodas vārdnīca*, was published in 1987 and has been reprinted several times.

Bilingual lexicography in Latvian is fairly advanced, with relatively big new dictionaries covering not only the traditional contact languages such as English, Russian, and German, but adding Estonian, Italian, Lithuanian, Polish, Spanish, and in the 1990s also Scandinavian and some Eastern languages. The last decade has seen the creation of new Latvian-English (Veisbergs, 2003), English-Latvian (2004), and French-Latvian (Auziņa, 2001) dictionaries as well as older German-Latvian and Russian-Latvian dictionaries reprinted and updated.

Among other important contemporary lexicographical sources one should mention dictionaries of etymology (Karulis, 1992), foreign words (Baldunčiks, 1999), idioms (Laua *et al.*, 1996), dialects, spelling, pronunciation, toponyms, abbreviations, numerous terminology dictionaries, and foreign name-spelling dictionaries (Latvian uses transcription for rendering foreign proper nouns).

Electronic Resources

Latvian lexicography suffers from the lack of a proper corpus. Other electronic tools are fairly well developed. A combination of various bilingual dictionaries is available from Tilde, and some of Tilde's resources are available on the Internet. An up-to-date and comprehensive Latvian explanatory dictionary with electronic access is being slowly built. Other portals include Valsts Valoda, the state language commission, and Valsts Valodas Aģentūra, the state language agency.

Various language-related resources, including Muehlenbach and Endzelins's dictionary, can be found on the website of the Artificial Intelligence Laboratory in the University of Latvia's Department of Mathematics and Computer Science. Termnet offers

Forssman B (2001). *Lettische grammatik*. Dettelbach: J. H. Röll.
Gāters A (1977). *Die lettische Sprache und Ihre Dialekte*. The Hague: Mouton.
Kabelka J (1982). *Baltų filologijos įvadas*. Vilnius: Mokslas.
Mathiassen T (1997). *A short grammar of Latvian*. Columbus, OH: Slavica.
Mīlenbahs K & Endzelīns J (1923–1932). *Latviešu valodas vārdnīca* (4 vols). Riga: Kultūras fonda izdevums.
Nau N (1998). *Latvian*. Munich and Newcastle: Lincom Europa.
Rūķe-Draviņa V (1977). *The standardization process in Latvian: 16th century to the present*. Stockholm: Almqvist & Wiksell International.

Latvian Lexicography

A Veisbergs, University of Latvia, Riga, Latvia

Latvian lexicography has been and remains primarily bilingual. This has been historically determined by the dominance of one or another country in the region – so that dictionaries accommodated to the situation accordingly. It is usual to date Latvian lexicography from 1638, when G. Mancelius's *Lettus* was published, containing about 6000 German-Latvian entries. There were also two multilingual dictionaries published in the same century: a Polish-Latin-Latvian work by Elger (1683) and a German-Latin-Polish-Latvian one by Dressel (1688). The following two centuries saw the same pattern – although the dictionaries increased in size and quality, they were made by non-Latvians. The situation changed in the middle of the 19th century, when after the national 'awakening' Latvian linguists took hold of lexicography. The new dictionaries were purged of German and Russian loanwords, and they came closer to spoken language yet also became more prescriptive, the Waldemahr (1872), with 50 000 entries. The dictionary scene up to 1900 was thoroughly researched by Zemzare (1961).

The iconic Latvian dictionary of the 20th century is the four-volume work by Muehlenbach and Endzelins, published in 1923, with supplements appearing in 1932 and 1946. It contains a wealth of Latvian word-stock, though purged of later loanwords. The dictionary contains roughly 120 000 entries in the new orthography with explanations (in German), etymologies, stress patterns, citations, and examples and, despite the passage of time and several shortcomings, it is still viewed as the 'proper' source of Latvian lexis.

The beginning of the 20th century saw an expansion of bilingual dictionaries in the new contact languages – English and French – as well as a strong emphasis on dictionaries of foreign words. The Soviet period produced the first extensive monolingual explanatory dictionary of literary Latvian (*Latviešu literārās valodas vārdnīca*) in eight volumes, which, however, suffers from a lack of colloquial language and is affected by ideological constraints. A one-volume desk dictionary (25 000 entries) of standard Latvian, *Latviešu valodas vārdnīca*, was published in 1987 and has been reprinted several times.

Bilingual lexicography in Latvian is fairly advanced, with relatively big new dictionaries covering not only the traditional contact languages such as English, Russian, and German, but adding Estonian, Italian, Lithuanian, Polish, Spanish, and in the 1990s also Scandinavian and some Eastern languages. The last decade has seen the creation of new Latvian-English (Veisbergs, 2003), English-Latvian (2004), and French-Latvian (Auziņa, 2001) dictionaries as well as older German-Latvian and Russian-Latvian dictionaries reprinted and updated.

Among other important contemporary lexicographical sources one should mention dictionaries of etymology (Karulis, 1992), foreign words (Baldunčiks, 1999), idioms (Laua *et al.*, 1996), dialects, spelling, pronunciation, toponyms, abbreviations, numerous terminology dictionaries, and foreign name-spelling dictionaries (Latvian uses transcription for rendering foreign proper nouns).

Electronic Resources

Latvian lexicography suffers from the lack of a proper corpus. Other electronic tools are fairly well developed. A combination of various bilingual dictionaries is available from Tilde, and some of Tilde's resources are available on the Internet. An up-to-date and comprehensive Latvian explanatory dictionary with electronic access is being slowly built. Other portals include Valsts Valoda, the state language commission, and Valsts Valodas Aģentūra, the state language agency.

Various language-related resources, including Muehlenbach and Endzelins's dictionary, can be found on the website of the Artificial Intelligence Laboratory in the University of Latvia's Department of Mathematics and Computer Science. Termnet offers

access to terminology containing various sources of language combinations from various times. The website of Tulkošanas un Terminoloģijas Centrs (Translation and Terminology Center) contains material dealing with EU matters.

Generally Latvian lexicography today bears a somewhat random character, but of late the descriptive approach seems to have prevailed.

See also: Bilingual Lexicography; German; Latvian; Lexicography: Overview; Phraseology; Russian; Terminology and Terminological Databases.

Bibliography

Auziņa I (2001). *Franču-Latviešu vārdnīca.* Riga: Zvaigzne.
Baldunčiks J (1999). *Svešvārdu vārdnīca.* Riga: Jumava.
English-Latvian Dictionary (2004). Riga: Jāņa Sēta.
Karulis K (1992). *Latviešu etimoloģijas vārdnīca.* Riga: Avots.
Latviešu literārās valodas vārdnīca (1972–1996). Riga: Zinātne.
Latviešu valodas vārdnīca (1987). Riga: Avots.
Laua A, Ezeriņa A & Veinberga S (1996). *Latviešu frazeoloģijas vārdnīca.* Riga: Avots.
Mancelius G (1638). *Lettus.*
Muehlenbach K & Endzelins J (1923–1946). *Latviešu valodas vārdnīca.* Riga.
Veisbergs A (2003). *The New Latvian-English Dictionary.* Riga: EAI.
Waldemahr K (1872). Russisch-Lettisch-Deutsche Wörterbuch. Moscow.
Zemzare D (1961). *Latviešu vārdnīcas.* Riga: ZAI.

Relevant Websites

http://www.ailab.lv – Artificial Intelligence Laboratory, Department of Mathematics and Computer Science, University of Latvia.
http://www.termnet.lv – Termnet.
http://www.tilde.lv – Tilde.
http://www.ttc.lv – Tulkošanas un Terminoloģijas Centrs.
http://www.valoda.lv – Valsts Valodas Aģentūra.
http://www.vvk.lv – Valsts Valoda.

Law and Language: Overview

D Kurzon, University of Haifa, Haifa, Israel

Introduction

Interest in legal language is not a modern phenomenon, although extensive research into it began to appear only in the latter part of the 20th century. Earlier comments on the language of law were exclusively critical of the style, vocabulary, and structure of legal texts, especially legislative language. We find the French philosopher Baron de Montesquieu in his *De l'Esprit des Lois* (1748) arguing for a simple legal language, to be understood by all readers in the same way, and furthermore to be understood by people even of mediocre understanding ("*de mediocre entendement*"). The 19th century English utilitarian, Jeremy Bentham, also proposed that legislation should be written in a language understood by laypersons. "Clearness in respect of its language" was one of the seven properties, Bentham argued, "a body of law . . . must be possessed of."

These comments show that there was a problem felt by leading thinkers concerning the language of the law. Both Montesquieu and Bentham related to legislation, but their criticism may be equally directed to other legal documents. Furthermore, legal language and legal institutions have been the target of criticism in literature. Charles Dickens' description of the English Chancery Court in his novel *Bleak House* – the time it takes for the judges to decide a case, the language which no one understands, etc., – is well known.

In the early days of modern linguistics, before it had become an established discipline, we do find studies on the law and its language. Several members of the pre-World War II Prague Linguistic Circle addressed problems of legal language. The language of legislators was examined from a functional perspective, with the focus on the educational level of readers of legislative texts, which implies that many people who are affected by the law are not able to read the texts in which the law is written. Another facet of the work of Prague linguists was in the field of vocabulary; observations were made on the careful use of words in legal texts which have inexact meanings in everyday discourse.

Academic lawyers had also dealt with legal language, and published articles, and even textbooks on the drafting of legislation and other documents. These tend to be prescriptive from a lawyer's point of view, and are not outwardly critical of the structural complexities found in legal texts. An exception is the seminal work of David Mellinkoff (Mellinkoff, 1963),

access to terminology containing various sources of language combinations from various times. The website of Tulkošanas un Terminoloģijas Centrs (Translation and Terminology Center) contains material dealing with EU matters.

Generally Latvian lexicography today bears a somewhat random character, but of late the descriptive approach seems to have prevailed.

See also: Bilingual Lexicography; German; Latvian; Lexicography: Overview; Phraseology; Russian; Terminology and Terminological Databases.

Bibliography

Auziņa I (2001). *Franču-Latviešu vārdnīca.* Riga: Zvaigzne.
Baldunčiks J (1999). *Svešvārdu vārdnīca.* Riga: Jumava.
English-Latvian Dictionary (2004). Riga: Jāņa Sēta.
Karulis K (1992). *Latviešu etimoloģijas vārdnīca.* Riga: Avots.
Latviešu literārās valodas vārdnīca (1972–1996). Riga: Zinātne.
Latviešu valodas vārdnīca (1987). Riga: Avots.
Laua A, Ezeriņa A & Veinberga S (1996). *Latviešu frazeoloģijas vārdnīca.* Riga: Avots.
Mancelius G (1638). *Lettus.*
Muehlenbach K & Endzelins J (1923–1946). *Latviešu valodas vārdnīca.* Riga.
Veisbergs A (2003). *The New Latvian-English Dictionary.* Riga: EAI.
Waldemahr K (1872). Russisch-Lettisch-Deutsche Wörterbuch. Moscow.
Zemzare D (1961). *Latviešu vārdnīcas.* Riga: ZAI.

Relevant Websites

http://www.ailab.lv – Artificial Intelligence Laboratory, Department of Mathematics and Computer Science, University of Latvia.
http://www.termnet.lv – Termnet.
http://www.tilde.lv – Tilde.
http://www.ttc.lv – Tulkošanas un Terminoloģijas Centrs.
http://www.valoda.lv – Valsts Valodas Aģentūra.
http://www.vvk.lv – Valsts Valoda.

Law and Language: Overview

D Kurzon, University of Haifa, Haifa, Israel

Introduction

Interest in legal language is not a modern phenomenon, although extensive research into it began to appear only in the latter part of the 20th century. Earlier comments on the language of law were exclusively critical of the style, vocabulary, and structure of legal texts, especially legislative language. We find the French philosopher Baron de Montesquieu in his *De l'Esprit des Lois* (1748) arguing for a simple legal language, to be understood by all readers in the same way, and furthermore to be understood by people even of mediocre understanding ("*de mediocre entendement*"). The 19th century English utilitarian, Jeremy Bentham, also proposed that legislation should be written in a language understood by laypersons. "Clearness in respect of its language" was one of the seven properties, Bentham argued, "a body of law . . . must be possessed of."

These comments show that there was a problem felt by leading thinkers concerning the language of the law. Both Montesquieu and Bentham related to legislation, but their criticism may be equally directed to other legal documents. Furthermore, legal language and legal institutions have been the target of criticism in literature. Charles Dickens' description of the English Chancery Court in his novel *Bleak House* – the time it takes for the judges to decide a case, the language which no one understands, etc., – is well known.

In the early days of modern linguistics, before it had become an established discipline, we do find studies on the law and its language. Several members of the pre-World War II Prague Linguistic Circle addressed problems of legal language. The language of legislators was examined from a functional perspective, with the focus on the educational level of readers of legislative texts, which implies that many people who are affected by the law are not able to read the texts in which the law is written. Another facet of the work of Prague linguists was in the field of vocabulary; observations were made on the careful use of words in legal texts which have inexact meanings in everyday discourse.

Academic lawyers had also dealt with legal language, and published articles, and even textbooks on the drafting of legislation and other documents. These tend to be prescriptive from a lawyer's point of view, and are not outwardly critical of the structural complexities found in legal texts. An exception is the seminal work of David Mellinkoff (Mellinkoff, 1963),

Language of the Law. In his book, Mellinkoff, a lawyer, is highly critical of the obtuse language lawyers employ in their documents.

Fields of Research

This section introduces the various fields of research that have grown over the past 30 years or so. Reference will also be made to the various articles that appear in this encyclopedia that relate to current research into those fields of law and language.

Linguistic interest in the language of the law was perhaps initiated by David Crystal and Derek Davy in their book on *Investigating English style* (Crystal and Davy, 1969). In this work, the authors examine a number of registers of English, including the legal register. In the chapter on legal language, they deal with the language of contracts, pointing out the style of legal draftsmen who have developed an intricate and syntactically complex style, in order, on the one hand, to prevent misunderstanding, and on the other, to cover all what the clients want to say and not to depend on extrinsic evidence for the interpretation of the document. They draw attention especially to (1) the conditional structure of the legal sentence; (2) the placing of modifiers immediately adjacent to the phrase or word to be modified; (3) the use of non-finite clauses as postmodifiers; and (4) the lack of explicit connection between the sentences of a document with its accompanying repetition of lexical items and the accompanying sparing use of pronouns. Following Crystal and Davy's book came a spate of work on legal language, especially the language of legal documents. A great deal of this work has been in the fields of grammar, especially syntax, and lexis. Much of the research was produced by Scandinavian scholars (e.g., Gustafsson, 1975), and several of the papers in the special issue of *Text* (1984) were devoted to the language of legal documents (*see* **Language of Legal Texts**) deals directly with written legal language, while other articles have a bearing on this register (*see* **Law on Language**).

In the United States, interest in the law and language came initially more from a sociological and anthropological than a linguistic perspective. One may also cite work on the language of Barotse jurisprudence, Yakan litigation, and Shawnee laws, among others. The type of doubletalk and gobbledygook used by government during the Vietnam War and especially during the Watergate affair also led to many researchers putting the focus on the language of administration and of law. Only in the early 1970s were sociolinguists introduced to this area of research. Brenda Danet's overview of language in the legal process, which appeared in the *Law and Society Review* in 1980, served as a stimulus to further and broader research in the areas of law and speech. Interest in oral communication in Western legal cultures grew, too, with the development of more efficient recording equipment, which allowed a more accurate record of spoken legal proceedings, thus allowing linguists and conversational analysts to tackle courtroom discourse. A conference was held in 1985 at Georgetown University on spoken legal discourse, e.g., witness questioning, depositions, jury instructions and interpretation; many of the papers presented were published by Levi and Graffam Walker (1990). Another research field in the area of law and spoken language addresses lawyer–layperson communication. This involves dialogue not only in the courtroom but also in the lawyer's office when the lawyer speaks to his/her client.

It can be seen that this distinction between written and spoken legal language is one that pervades research into legal language. Taxonomies of types of legal language have been set up. Within the mode of written legal language, not only are documents such as legislation and contracts examined, but also legal textbooks and law reports. Spoken legal language may be subcategorized into courtroom discourse and out-of-court discourse, and each of these subcategories may be further subdivided; see a description of genres in general (*see* **Legal Genres**).

Against the background of growing interest in spoken legal language, the study of written legal language may seem to have reached its peak in the 1980s; further work on legislative texts, on contracts, and on wills has been carried out not so much from a structural point of view, but more from a discourse and pragmatic perspective. Legal texts, especially written legal documents, have been analyzed as discourse. Moreover, pragmatics has been extended to speech in legal proceedings, for example the adjacency pair of question/answer in witness testimony, and formulaic language in texts such as the witness's oath. This type of spoken legal language has been, and is still being, extensively studied within a sociolinguistic framework, in which pragmatics has found an important niche. The reader is directed to the articles on legal pragmatics and on definitions and rules in legal language.

At the same time as these developments, another field opened up, a field that has emerged from time to time since Montesquieu and Bentham, but has not until recently attracted the attention of scholars in linguistics and sociolinguistics – the problem of the readability and comprehensibility not only of legal texts but, principally, of spoken legal discourse. The Charrows (1979) conducted a psycholinguistic study of jury instructions, pointing out the low level of

understanding of many jurors, who may have to make fateful decisions concerning matters they do not grasp. Many of the attempts to improve the style of legal documents, to make them comprehensible to the layperson, derive from work on language structure. This research and accompanying public agitation eventually led to the support by U.S. President Carter for plain language in legal documents that directly affect the public, e.g., insurance policies. Many legal firms, especially but not only in the English-speaking world, have adopted plain English in their document writing as company policy (*see* **Jury Instructions**).

The right of silence of the accused, especially in the Anglo–American legal context, has been studied both in terms of its pragmatics, with attempts made to answer the question what the accused means by his or her silence during police investigation, and also in terms of comprehensibility, establishing it as a field of research associated with plain language. Not only has the Miranda warning (in the United States) been under scrutiny as far as comprehensibility is concerned, but in the wake of changes in the text setting out the accused's right of silence carried out in the United Kingdom, linguists have investigated whether the average suspect with a below-average IQ may understand what is being read out to him or her, and what s/he is reading.

Apart from suspects who may not understand what is being said to them by authoritative figures (e.g., police and lawyers), there are other disadvantaged groups who face a linguistic battle in the course of legal proceedings. Litigants and suspects who do not speak, or who have an inadequate knowledge of, the language used in the courts are in many jurisdictions entitled to an interpreter. The various problems arising from interpretation and the rights of the interpreter and the interpreted are discussed in **Court Interpreting**, while other disadvantaged groups, for example aborigines in Australia, which has attracted much research recently, is dealt with in **Interviewing and Examining Vulnerable Witnesses**.

The employment of linguists as expert witnesses in the courtroom has grown over the last two decades or so. The expertise work requested tends to be in the field of voice and writing identification. Much of the research on language in the courtroom is now included in this field of studies. The associated field of intellectual property rights and their violation is an area of current research discussed in **Plagiarism**.

We also find groups of philosophers, linguists, and academic lawyers who examine law and language from a semiotic point of view. The French semiotician Andre Greimas wrote with linguist Eric Landowski a seminal work on legal discourse (1966), followed by further work from a Greimasian perspective by Bernard Jackson (1984). Semiotic research into legal language has also appeared from a Peircian standpoint, and from the Italian analytical school of jurisprudence. A discussion of police/citizen encounters from a semiotic point of view is found in **Legal Semiotics** and **Police Questioning**.

The development of legal language has been studied, especially within the European area. Despite a distinction that needs to be made between common law (Anglo–American) law countries and countries whose legal systems are what is termed continental, developed from the Justinian code to the Napoleonic code and beyond, the language of the law in both systems has much in common. This may be seen, too, in a comparative approach in non-European systems. Furthermore, studies have been carried out on the influence of the law on 'normal' language itself, as may be seen in **Law on Language**.

There remain a number of less prominent areas in which research has been carried out on the law and language. For example, within the model of Roman Jakobson's functions of language, the poetic function of the language of the law has been examined in terms of rhythm, alliteration, parallel structures, and the extensive use of metaphor.

See also: Conversation Analysis; Court Interpreting; Interviewing and Examining Vulnerable Witnesses; Jury Instructions; Language of Legal Texts; Law on Language; Legal Genres; Legal Semiotics; Plagiarism; Police Questioning; Prague School; Silence; Understanding Spoken Discourse.

Bibliography

Bhatia V (1983). *An applied discourse of English legislative writing.* Birmingham: University of Aston.

Charrow R P & Charrow V R (1979). 'Making legal language understandable: a psycholinguistic study of jury instructions.' *Columbia Law Review 79,* 7.

Crystal D & Davy D (1969). *Investigating English style.* London: Longman.

Danet B (1980). 'Language in the legal process.' *Law and Society Review 14(3),* 445–564.

Danet B (1984). Special issue on legal language. *Text 3,* 1–4.

Forensic linguistics: the International Journal of Speech, Language and the Law.

Gibbons J (ed.) (1994). *Language and the law.* London: Longman.

Greimas A J & Landowski E (1979). *Introduction à l'analyse du discours en sciences socials.* Paris: Hachette.

Gumperz J & Hymes D (eds.) (1972). *Directions in sociolinguistics.* New York: Holt, Rinehart and Winston.

Gustafsson M (1975). *Some syntactic properties of English law language.* Publication of Department of Türkü. 4.

International Journal for the Semiotics of Law.

Jackson B (1997). *Semiotics and Legal Theory* (2nd edn.). Liverpool: Deborah Charles Publications.
Kurzon D (1986). *It is hereby performed . . . explorations in legal speech acts*. Amsterdam: John Benjamins.
Kurzon D (1998). *Discourse of silence*. Amsterdam: John Benjamins.
Levi J N & GraffamWalker A (eds.) (1990). *Language in the judicial process*. New York: Plenum Press.
Mellinkoff D (1963). *Language of the law*. Boston: Little, Brown.
Shuy R W (1993). *Language crimes: the use and abuse of language evidence in the courtroom*. Oxford: Blackwell.
Tiersma P M (1999). *Legal language*. Chicago: Chicago University Press.

Law on Language

A Wagner, University of the Littoral Côte d'Opale, Boulogne sur mer, France

Contributions of the History of English Civilization and Its Influence on Language and Discourses

The exploration of modes of communication in Great Britain permits a better understanding of the mechanisms by which legal order has been integrated into everyday language. The English judicial system has evolved since the 13th century under the influence of the customs of the different ethnic groups that have populated the island. Sir Francis Bacon noted the multicultural aspect of the Laws of England, the contents of which were the result of a compromise between the natural order of things and tribal or clannish values and the desire to homogenize through organized constraints.

The current language system is the result of an organization of meanings and terms from different cultures and traditions imported by invaders. They contributed to the impression and development of oral language and the transmission of orders and knowledge. What is the influence of the various ethnic groups that have contributed to the emergence of the nation?

The Celts, Anglo-Saxons, and Scandinavians participated in the historical construction of the linguistic dialectic between law and language in England. By successive invasions, they transposed their traditions through the development of rules for social life. At that time, customary practices were systematically built because illiteracy was virtually universal: "It is true, indeed, that, in the profound ignorance of letters which formerly overspread the whole western world, all laws were entirely traditional, for this plain reason, that the nations among which they prevailed had but little idea of writing" (Blackstone in Mellinkoff, 1963: 158).

This "plain reason" had fundamental consequences for the creation of an English legal order, which over the centuries came to be based on customs because of the need to articulate the practices and beliefs associated with each ethnic group; the purported objective of such a move was the creation of national unity. Writing was rare, but where it existed, it acquired a supreme value through its solemnity; the Magna Carta of 1215 is a good example, for it established, in English law, the mechanism of the predominantly customary parliamentary system that existed at the time. King Charles I received the ultimate punishment in 1649 for attempting to defy Parliament and trying to modify the Presbyterian religious custom.

The Celts

The influence of the Celts, who practiced human sacrifice and vengeance, bestowed a precise, very complex customary law through their elaborate system of sanctions. The lawyers were artisans of a repetitive oral tradition: "The joint memory of the ancients, the transmission from one ear to another, the chanting of the poets" (Powell, 1985: 159).

Sacred centers (*omphalos*) assured the unity between customs – beliefs and the expression of common language. Notably, the Celts were at the origin of 'legal procedures.' Even if justice was individual, it was rendered by an assembly, meeting as a jury. Resolving conflicts by arbitration resulted in making the law more fluid.

The Justinian Institutes

The Justinian Institutes were the source of judicial language carried by the Romans into the towns they occupied. These were the only hearths of latinization (Walter, 1994: 366) in an environment dominated by the Celtic spirit. The contributions were substantial, as the Latin society imposed the tradition of writing by the intermediary of the clergy, by which the

Jackson B (1997). *Semiotics and Legal Theory* (2nd edn.). Liverpool: Deborah Charles Publications.
Kurzon D (1986). *It is hereby performed . . . explorations in legal speech acts*. Amsterdam: John Benjamins.
Kurzon D (1998). *Discourse of silence*. Amsterdam: John Benjamins.
Levi J N & GraffamWalker A (eds.) (1990). *Language in the judicial process*. New York: Plenum Press.
Mellinkoff D (1963). *Language of the law*. Boston: Little, Brown.
Shuy R W (1993). *Language crimes: the use and abuse of language evidence in the courtroom*. Oxford: Blackwell.
Tiersma P M (1999). *Legal language*. Chicago: Chicago University Press.

Law on Language

A Wagner, University of the Littoral Côte d'Opale, Boulogne sur mer, France

Contributions of the History of English Civilization and Its Influence on Language and Discourses

The exploration of modes of communication in Great Britain permits a better understanding of the mechanisms by which legal order has been integrated into everyday language. The English judicial system has evolved since the 13th century under the influence of the customs of the different ethnic groups that have populated the island. Sir Francis Bacon noted the multicultural aspect of the Laws of England, the contents of which were the result of a compromise between the natural order of things and tribal or clannish values and the desire to homogenize through organized constraints.

The current language system is the result of an organization of meanings and terms from different cultures and traditions imported by invaders. They contributed to the impression and development of oral language and the transmission of orders and knowledge. What is the influence of the various ethnic groups that have contributed to the emergence of the nation?

The Celts, Anglo-Saxons, and Scandinavians participated in the historical construction of the linguistic dialectic between law and language in England. By successive invasions, they transposed their traditions through the development of rules for social life. At that time, customary practices were systematically built because illiteracy was virtually universal: "It is true, indeed, that, in the profound ignorance of letters which formerly overspread the whole western world, all laws were entirely traditional, for this plain reason, that the nations among which they prevailed had but little idea of writing" (Blackstone in Mellinkoff, 1963: 158).

This "plain reason" had fundamental consequences for the creation of an English legal order, which over the centuries came to be based on customs because of the need to articulate the practices and beliefs associated with each ethnic group; the purported objective of such a move was the creation of national unity. Writing was rare, but where it existed, it acquired a supreme value through its solemnity; the Magna Carta of 1215 is a good example, for it established, in English law, the mechanism of the predominantly customary parliamentary system that existed at the time. King Charles I received the ultimate punishment in 1649 for attempting to defy Parliament and trying to modify the Presbyterian religious custom.

The Celts

The influence of the Celts, who practiced human sacrifice and vengeance, bestowed a precise, very complex customary law through their elaborate system of sanctions. The lawyers were artisans of a repetitive oral tradition: "The joint memory of the ancients, the transmission from one ear to another, the chanting of the poets" (Powell, 1985: 159).

Sacred centers (*omphalos*) assured the unity between customs – beliefs and the expression of common language. Notably, the Celts were at the origin of 'legal procedures.' Even if justice was individual, it was rendered by an assembly, meeting as a jury. Resolving conflicts by arbitration resulted in making the law more fluid.

The Justinian Institutes

The Justinian Institutes were the source of judicial language carried by the Romans into the towns they occupied. These were the only hearths of latinization (Walter, 1994: 366) in an environment dominated by the Celtic spirit. The contributions were substantial, as the Latin society imposed the tradition of writing by the intermediary of the clergy, by which the

imprint of Christianity on the discourse was complete. These written acts made knowledge accessible even with a limited level of literacy. A transformation of the transmission of knowledge and its fixture by written documents can be observed in the numerous royal edicts issued. However, this mutation only concerned an elite. From then on, Latin terms carried both religious and legal meaning. For example, on *Coelesti* and *Inferi*:

> Christianity gave a new meaning to both words. Hell became a place of torment, inhabited by devils, oath-breakers, and other condemned souls. And it was now understood that a good man might live forever in heaven. This shift of emphasis put considerably more wallop into the oath and a deed or will drawn by a literate clergyman who often threatened hell-fire and damnation for those who tried to upset the transfers (Mellinkoff, 1963: 63).

In Latin terminology, for example, the word *cell* (latin: *cella*, Old French: *celle*) meant both a cell in a convent or monastery but would also come to mean a prison.

The Anglo-Saxons

The Anglo-Saxons founded their rules of community life on the principles of family vengeance and redress: *the Wergeld* (compensation). The law of Ethelbert of Kent concerning the social and matrimonial system is an excellent example. The laws enforceable in Kent in effect transposed the Germanic order to protect women, widows, and slaves. On the other hand, they treated murder and theft very harshly (Stenton, 1968 : 97–98).

The legacy of Old English (Mellinkoff, 1963: 46–47), permitting the creation of words by composition, has come down to modern English, and is still used today:

- boc — book, charter, deed
- bocland — land governed by the charter
- folcland — land not governed by the special terms of a charter, but by folk custom
- wer — a man (sometimes also used to mean wergild)
- wergild — man compensation
- wergildpeof — a thief whose criminality had been ended by the payment of his wergild

In modern texts, the terms *hereafter, herein, hereof, herewith* are occasionally found. This historical contribution of Anglo-Saxon adds to the common language by making the terms more precise or even distorting them.

It is one of the first steps towards the creation of a specialized language giving birth to specially adapted expressions. (Gény, 1922, T.3: 460.)

The Normans

The Norman influence reveals the catalyzing effect of written Latin then in use in France. It became one of the principal instruments of diffusion of English law, notably in the elaboration of rules of law and charters. William the Conqueror helped perpetuate the use of Latin because he had understood the limitations of Old English and Old French for expressing Norman judicial ideas (Holdsworth, 1923: 44). Latin was the language *par excellence* of the medieval age (Woodbine, 1943: 411). The concept of Law Latin is described in no uncertain terms: "It is 'barbarous,' 'corrupt,' 'mutilated,' 'dog latin,' and in an Irish version 'bog Latin.' The calmest definition is Webster's: the kind of Low Latin, containing Latinized English and old French words, used in English law" (Mellinkoff, 1963: 73).

William the Conqueror adapted Norman judicial institutions to England, separating the ecclesiastical courts from the secular, but also continuing to use clergymen in judicial posts, thereby perpetuating the use of Latin. He affirmed the compromise in 1086 when he called a grand assembly at Salisbury to: "… cement the union between the indigenous and Norman elements [and to] give the structure an unchallengeable judicial foundation" (Andrieu-Guitrancourt, 1978: 278).

While the English were establishing a power to regulate the life of society, the French were also attempting to unify their judicial practices under French customs. The procedure was the same as that on the other side of the Channel. The '*decretale*' of Innocent III in 1202 addressed to Lothaire, Canon of Paris, posed the problem of Parisian customs (*Coutume de Paris*) directly and incisively. He referred to this custom as a part of the law in France (*quia tamen juxta, consuetudinen approbatam, que pro lege servatur, in partibus gallicanis.*) A great number of other 13th century charters relating to it were issued as law in 1212 by Simon de Montfort to the people of Albigeois and other neighboring regions (charters following the Treaty of Franc-Aleu of 1637).

It is also noticeable that insularity did not drive England to draft original rules parallel to the expression of a multiethnic population. The construction of a derived specialized language is the consequence of a complex maturation of Norman and foreign customs. It results in a common law highlighted by the clergy who integrated Canon law as well as Roman law into it.

The Setting of 'Open Texture'

An 'open texture' (MacCormick, 1978) leads to various possibilities of interpretation within space and time. Indeed, the interpretation of the law is currently made in light of its principal sources, which are statutory law and judicial precedents. The written word is of supreme authority, but is integrated in the English logic of interpretive reasoning. In contrast to countries based on written law where the concept does not allow any transformation, its use in England is flexible. This flexibility is proof that the capacity of the legislative discourse, which is situated – by Act – outside of time and space, can become a living reality. Butler (1997: 25) defines its sense when she affirms that the language "remains alive when it refuses to 'encapsulate' or 'capture' the events and lives it describes." This doctrine is aimed at and authorizes the interpretation of discourse. The two elements that shape the decision are the power of deformation of concepts associated with vague outlines in discourse. Their use by the legislator is not accidental: it is the culmination of a premeditated act whose purpose is to simplify the formulation of written law. Indeterminacy is characteristic of the application of judicial precedents:

> In language the words which present themselves, and are employed in the character of names, are, some of them, names of real entities, others, names of fictitious entities; and to one or other of these classes may all words which are employed in the character of names be referred (Bentham in Tait, 1843: 448–449).

Indeterminacy can also be noticed in the gender, the words, the sense accorded to concepts, and the interpretations in time and space. It reveals the specific identity of its legal and linguistic communities: what is the approach of this interpretation?

Fluidity of Meaning: Cause or Consequence of Judicial Vagueness?

In law, vagueness is understood to be the remnant of doubt in the legal text, outside of the clear meaning flowing from its pronouncement. Plurality in meaning is advantageous since the possibility of more messages implies a larger set to select from.

> But now I employ the expression 'Legal Fiction' to signify any assumption which conceals, or affects to conceal, the fact that a rule of law has undergone alteration, its letter remaining unchanged, its operation being modified... The fact is [...] that the law has been wholly changed; the fiction is that it remains what it always was (Sumner Maine, 1963/1861: 25–26).

This fiction, this voluntary lack of real frontiers between the semiotic logic and the multiplicity of the fields of application leads to instability in discourse between scientific communities. But these 'neighborhood' conflicts are considered beneficial to legal language because of the important role left for interpretation. The English philosopher, John Heron (1989), confronts the sense of this collaboration:

> This construing of what is immediately presented as a spatio-temporal whole that transcends what is immediately presented, is not only a means of experiential knowing, it is also a fundamental kind of practical knowledge – knowing how to orientate oneself in space and time, knowing how to construe presentations or appearance in terms of spatial and temporal 'reality.'

This space for creative liberty is not the result of a misunderstanding or of a lack of completion. It is based on the semantic contents of the text, which by its organization, inserts it into a general body of reflection: the doctrine of indeterminacy. This particular form, for example, places the judge in a relatively open space for decisions; it offers him the possibility to integrate exterior elements into the legal domain within prescriptive arguments. Indeterminacy transforms the partially hermetic reading of the judge who applies the law, in judgments, into a flexible interpretation – *ratio decidendi.*

This doctrine produces social identity but is also a source of cohesion between the individuals of the same scientific community. Bachelier reminds us that "the richness of a scientific concept is measured by its power of deformation" (Grawitz, 1993: 1). This flexibility of concept becomes a living reality, "a kind of **personification** ... as if the law were a living creature" (Bentham in Tait, 1843: 399). The rigid prescriptive frame delimits a content of which the sense varies according to the interpretation of the user.

Theoretical Intepretation The theoretical interpretation of the author's thoughts and his/her productions permits the discernment of several essential elements that condition the tenor of the messages. There is no perfect accord between the thought of the author and his/her production. The deformation operates from the relationship of factual or written elements, which serve to support his/her thoughts. The ambiguity comes out of the personal work and goes on to contribute to the specificity in the identity of the discourse. The elements of shift in meaning are therefore either in the semantic support between constant or variable meaning, or in a time-space that does not correspond to that of the emission of the message. The networking of the listed constraints makes sense and leads to the determination of rigid and fragile elements (i.e., fuzzy sets in the language of the

law that can impair the whole understanding of a sentence). It may be pointed out that writing down the rule is the product of a semantic reflection in which textual organization also takes on sense.

Metacognition The metacognition of the interpreter is important because of the effects of perception of ambiguity integrated in the author's discourse. To that cultural appreciation is added the special knowledge of the reader who has been trained in legal logic. Flavell (1987: 21–30) underlines the effect of doubt, which is present in the cognitive activities that one does not know how to interpret or explain, and in which the construction of knowledge is proportional to the metacognitive implication of the subject.

Application of Ambiguity The application of ambiguity in the judicial act is realized in the form of a readjustment including the user's reasoning through a descriptive process. These contributions are characterized by the emission of procedural rules, drafted in part from law reports conditioned by the rules of precedent. The resulting process provides a compromise between the social parameters introduced and the decision reached. Its aim is to be in a continuum of the legal thought of the time: "Any instance of interpretation always involves a choice and a motive, an element of the 'play of meanings' which denies that there can ever be a single 'correct' or valid interpretation (Goodrich, 1986: 138)."

This handling of meaning and words is the prime feature of the functioning mode of the English legal system. The contribution of specialized linguistics allows the evaluation of the level of logic and fluidity that impregnate the frames of statutory law and precedents, that is of the drafters, readers, and interpreters of the law. This creation of linguistic tools is a challenge to the objectivity of the rules. Some writers mention the limitations made necessary by the fact that realities of social life only become judicial realities thanks to the concepts and decision making capable of providing the necessary meaning. The decision making outside the context of the rule interferes with the legal order but allows the extension of meaning, which can be re-used in other circumstances. This plural reality of linguistics can be noticed in the thinking of Sherry (1996: 134): "The globalization of culture means that we all live in 'translated' worlds that the spaces of knowledge we inhabit assemble ideas and styles of multiple origins, that transnational communications and frequent migrations make every cultural site a crossroads and a meeting place."

Gény (1922: 79–80) completed his notion of fluidity by a more global approach:

> The things of the world submitted to duration are in a constant flux and as the object of a continuous flow in such a way as there is no fixed nature but essences perpetually changing and observing at the same time that the faculties of man do not operate according to a law of schematical dissociation, they operate in a perpetual concert which does not allow to insulate them from one another. The 'new philosophy' energetically criticizes as means of knowledge, the 'discursive process' which leisurely deform reality (like the steps of a stairway which would be substituted to a slow slope), even the pure concepts which only grasp things by fragments and represents them like still and solid moulds incapable of expressing their essential variety and fluidity.

This brief analysis can be summarized by the fact that fluidity is a place of contact, a place of interference, and a mode of cultural production, 'a translated culture' as part of the processes by which newness enters the world (Wagner, 1999).

Indeterminacy in Legal Language

As an illustration, the current language runs on a path of information by means of expression neutral and accessible to everybody. Specialized languages run on the same path but add specialized elements to the current meaning. It, in turn, creates a dynamic of comprehension, a metalanguage only accessible to the initiated. The question is to examine the origin of the specialized term and its semantic transformation.

In the domain specific to legal expression, the linguist works at confronting the conceptual word within the current relationship and the application being construed by the powers that be. He or she then attempts to explain the mechanisms that condition the understanding of meanings both internal and external to the British scientific community. The distortion of meaning of this affected language is particularly perceptible by the other inhabitants of the European continent.

So, the question of reassigning meaning to words and playing with words will continue to be the core issue in legal adjudication, since any word may be subject to interpretation as it was with the case of 'tomatoes' in 1893 (*Nixen v. Hedden*) where the U.S. Supreme Court had to decide whether they were fruit or vegetables. The Court rejected the botanical truth that the tomato is in fact a big berry, and deferred to the culinary vernacular of vegetable to describe it.

As a consequence, statutory wording is not a closed system but an open texture where relative

indeterminacy is used to interpret statutes (Wagner, 2005). "Parliament often uses general words intentionally in the belief that judges are the most adept in filling in the details" (Freeman, 1994: 193).

Indeterminacy will only cease when one momentary experience (i.e., a case in court) is considered. It will then turn into a contextualized determinacy. "Because the meaning of language is inherently contextual, we have declined to deem a statute 'ambiguous', for purposes of lenity merely because it was possible to articulate a construction more narrow than that urged by the Government" (Solan, 2004: 79).

Involuntary Intent of Promoting Ambiguity The language contains several microsystems, which merit examination. Here, the difficulty of interpreting 'and' and 'or' is a well-known phenomenon in law. In a statute, these two terms can be interchangeable from the time that a legislative intention exists (Solan, 1993: 45). Quirk (1986: 934) in the same vein, adds "*or* tends to replace *and* in contexts which we have called non assertive, and more generally in subordinate positions in the sentence." Solan explained that an error in writing by the overuse of the terms 'and' and 'or' allows to invert them. In effect, they are considered by lawyers to be very vague, therefore too general. Their usage as indeterminate notions reflects the legal terminology. In this way, the weak interpretive use of the connector forces the court to take the necessary measures to arrive at the legislative intention. Mellinkoff (1965: 56) said the same thing: "Legal writings continue to spew contradictions joined by a habitual *and/or.*"

This controversy is very important to the use of this type of expression. Can it then be considered as an encouragement of mental laziness? Anyone knowing the difference between 'and' and 'or' can take a contract or a law and substitute one or the other term to reach a greater clarification of the law or legal decision.

The term 'and' is used to avoid a construction which, by the use of 'or' only, excludes the combination of several alternatives. 'And' takes the place of an addition, after several connected alternatives, simply by 'any combination thereof.' It takes the place of 'or' and in keeping the same meaning in statutes: 'and any combination thereof' or 'or any combination thereof.' By this, 'and/or' is completely ambiguous. But the courts will accept it without difficulty as a contribution to legal English: "The high failure rate of legal papers that depend on *and/or* for anything of importance should have long since eliminated it from the legal vocabulary. Yet it persists, and is widely used in legal and ordinary writings" (Mellinkoff, 1965: 55).

'Promoted' Indeterminacy Overtly 'promoted' indeterminacy in statutory law was emphasized by U.K. legislators on the use of gender-neutral, gender free language in order to avoid designating grammatical gender and leave doors open to judges' construction. The English Interpretation Act 1978 provides that:

"words importing the masculine gender include the feminine";
"words importing the feminine gender include the masculine," and
"words in the singular include the plural and words in the plural include the singular" (Wagner, 1999).

Even modality has been construed by judges as shown below:

> It is admitted that the words 'every member . . . shall inform' the directors does create an obligation but it is argued by the defendants that the words 'the directors . . . will take the shares' imports in some way the idea of an option or choice or volition on the part of the directors having regard to the inherent difference (not always observed) in the English language between the words 'shall' and 'will.' I appreciate the force of that argument, but I cannot accept it. In this context, while the word 'shall' clearly imports compulsion and obligation, the word 'will' indicates as it seems to me a resultant prospective eventuality, in which the member has to sell his shares and the directors have to buy them, each being under an obligation to bring that eventuality into effect (*Rayfield v Hands* [1958] 2 All ER 194 at 196, per Vaisey J).

> We must endeavour to construe the words, so that "shall and may" [in an agreement for arbitration] may both stand. "Shall and may" will both stand, by our holding them to be imperative. (*Crump v Adney* (1833) 1 Cr & M 355 at 361, per Lord Lyndhurst CB.)

> The words shall and may in general acts of parliament, or in private constitutions, are to be construed imperatively. (*A-G v Lock* (1744) 3 Atk 164 at 166, per Lord Hardwicke LC.)

This deliberate intention to promote indeterminacy is also visible when using adjectives or adverbs that could lead to an interpretive process from the judge. Following the examination of distinctive statutes in various countries, it can be noticed that using such words allows focus on intentional possibilities of interpretation. As an example, under the English Judicial Trustees Act 1896, s.3(1), plurality in interpretation is suggested when the drafters inserted adverbs such as 'reasonably' or 'honestly,' which

convey an emotional content. Indeed, both adverbs are vectors of perception of the universe, and from one case to another interpretation can easily fluctuate. Another example is to be found under the proviso of the Australian Criminal Law Amendment Act 1922, s.2, where it is provided that there is a "reasonable cause to believe that a girl is of or above the age of 16 years." How do we determine this reasonable cause? Only a specific case can give an answer to the question.

The Canadian Landlord and Tenant (Rent Control) Act 1949, s.3 (1) (b) deals with the 'reasonable price' to be charged. Here, the evaluation of 'reasonable price' is highly dependent on the context. In other words, indeterminacy is a medium of communication permitting the nondetermination of the constituent words. Words lack effective referents. They are idle, they are not connected with a 'living reality' but with statutes, and their determination will be provided only when exemplified.

Even the simplest words ('person' and 'night' in this case) can have an abundance of meanings in the light of distinctive Acts, for example:

- In the Statute of Frauds Amendment Act 1828, it is provided that 'person' includes a 'corporation.'
- Under the Interpretation Act 1978, Sch.1, 'person' includes a body of persons corporate or unincorporated.
- In the light of Railway Clauses Consolidation Act 1845 (c.20), s.20 'person' includes a limited company.

- What time shall be considered 'night': provided always that for the purposes of this Act the night shall be considered and is hereby declared to commence at the expiration of the first hour of sunset and to conclude at the beginning of the last hour before sunrise (Night Poaching Act 1828, s.12)
- 'Night' "Means the time between half an hour after sunset and half an hour before sunrise" (Highways Act 1980, s.329(1)).

Hypercodification of Legal Language

Hypercodification is one of the characteristics of modern societies, which want to embrace the multiple facets of social life by speech or writing. One of the defining features of modern societies is their embrace of plural writing and speaking systems, a trend that is mirrored in the diversity of social life in general. It is a quest of meaning, the modalities of which seem to pervert the spirit. The exploration of this state of facts does not limit itself to its legal signification but circles the subject in its linguistic expression. This overabundance of modeling systems disturbs the conditions of emission, reception, and meaning of the messages.

The semiotic approach is one of the unavoidable modalities where investigating is carried out by the analysis of the practices of the legal community, notably in its system of values and behavior.

The hypercodification, which is very important in English Law, is part of the message. It can be understood as a means, a mode of expression showing that the legal culture is being communicated. The understanding of the phenomenon results from the exceptional conditioning of the discourse by historical heritage, the multiform contributions being the result of the mixing of cultures and customs of a multiethnic nation. Besides, the external interference has led to the amplification of legal productions and processes. Culler (1981) describes the mechanisms, which have resulted in the inflation of the legal system by the internal interrelations between written and symbolic languages.

Linguistic Territory

The linguistic territory has, in its organization, multiple forms of actions on law by the effects of symbolization, codification, normalization, and even standardization of the language. All of the operations undertaken are intended to simplify the structure of writing and make it more comprehensible. The main idea is to limit the length as well as the complexity of phrases. To achieve this simplification, several techniques have been elaborated. The stated goal is to put into practice a codification that is relatively stable, simple to understand and capable of being added to. In effect, the realization puts in place the means which, under the pretext of simplification, contribute to the complication of the original sense of the text.

Legal Identity

Legal identity is under constant reconstruction. In reality, this reconstruction is a legal one, which adds to the already existing foundation. This new formulation leads to modifications in the discourse by means of textual contributions, which create more inconveniences than advantages. The encoding entails an overabundance of codification. Its interpretations become more and more complex over time. They create an insidious rigidity of legal language. The consistency of discourse is dependent upon its comprehension; so as it is meant first of all for initiates, people outside of the community cannot understand it. The heritage of the Common Law is rich in concepts, in particular meanings, which hold true

only for England and Wales. To show how the law has become a medium, a vehicle, in the same way as the common language, of a highly codified and specialized language, it is necessary to explore its terminological construction. This necessity results from structured collective procedures by systems of abbreviation and terms of variable content. Hypercodification is so highly visible in legal language and the understanding of legal language varies within space and time.

A Tool of Socialization

The language of the law as a tool of socialization and factor of identity generates subproductions, which add new rules by its willingness to adapt in time and space to the English population.

The conveyed message traps legal language by the implication of networks and persons in which specific codes are structured in socio-autonomous systems. One example among others is in English medical law (The Human Fertilisation and Embryology Act 1990), in the case of surrogate mothers, where the adoptive mother is called the 'genetic mother.' In society, the mother who delivers the child is by definition the genetic mother.

Conclusion

Law is, by definition, the codification of modes of expression in a given society. However, its use is distinctive owing to the circumstances in space and time in which users apply or modify it. The shift in meaning between the product and its expression has been analyzed showing how the language of the law can be multifaceted.

See also: Law and Language: Overview; Legal Genres.

Bibliography

A-G v Lock (1744) 3 Atk 164 at 166.

Andrien-Guitrancourt P (1978). *Histoire de l'Empire Normand*. Paris: Payot.

Butler J (1997). *Excitable speech: a politics of the performative*. London: Routledge.

Crump v Adney (1833) 1 Cr & M 355 at 361.

Culler J (1981). *The pursuit of signs: semiotics, literature, deconstruction*. London: Routledge & Kegan Paul.

Flavell J H (1987). 'Metacognition, Motivation and Understanding.' *Revue de Metaphysique 2,* 21–30.

Freeman M D A (1994). *Lloyd's introduction to jurisprudence*. London: Sweet & Maxwell.

Gallard (ed.) (1637). *Chartes en suite du traité du Franc-Aleu*. Paris: Gallard Edition.

Gény F (1922). *Science et technique en droit privé positif.* (3 Tomes) Paris: Recueil Sirey.

Goodrich P (1986). *Legal discourse: studies in linguistics, rhetoric and legal analysis*. New York: St Martin's Press.

Grawitz M (1993). *Méthodes des sciences sociales*. Paris: Dalloz.

Heron J (1989). *Philosophical basis for a new paradigm*. London: Plenum Press.

Holdsworth L (1923). *A history of English law.* London: Sweet & Maswell.

MacCormick N (1978). *Legal reasoning and legal theory.* Oxford: Oxford University Press.

Mandl H, Stein N L & Trabassa T (eds.) (1968). *Learning and comprehension of text*. New Jersey: Lawrence Erlbaum Associates.

Mellinkoff D (1963). *The Language of the law.* London: Pluman.

Mellinkoff D (1965). *Legal writing: sense and nonsense*. London: Plunam.

Nixen v. Hedden, 149 U.S.304 (1893).

Pasquier M Le Duc (1847). *L'interprétation des Institutes de Justinien d'Etienne Pasquier 1529–1615*. Paris: Videcocq et Durand.

Powell T G E (1985). *The Celts*. London: Thames and Hudson.

Quirk R (1986). *Words at work: lectures on textual structure*. London: Longman.

Rayfield v Hands [1958]. 2 All ER 194 at 196, per Vaisey J.

Sherry S (1996). *Gender in translation*. London: Routledge.

Solan L (1993). *The language of judges*. Chicago: Chicago University Press.

Solan L (2004). 'Philosophical, psychological, linguistic, and biological perspective on legal scholarship I: Law and language: pernicious ambiguity in contracts and statutes.' *Chicago-Kent Law Review 79,* 859.

Stenton F M (1968). *Anglo-Saxon England*. Oxford: Clarendon Press.

Sumner Maine H (1963/1861). *Ancient Law 24*. London: Beacon Press.

Tait W (1843). *Jeremy Bentham Works 73*. Edinburgh: John Bowring.

Wagner A (1999). 'La validité de la normalisation legislative anglaise.' *International Journal for the Semiotics of Law 12(1),* 3–26.

Wagner A (2002). *La langue de la Common Law.* Paris: L'Harmattan.

Wagner A (2004). 'Le diagnostic de la pluralité en droit médical anglais et ses implications.' *Semiotica 151(1/4),* 183–200.

Wagner A (2005). 'Hypercodification de la Common Law: en quête de sens.' *Hermes 34*, 245–263.

Wagner A (2005). 'Semiotic analysis of the multistage dynamic at the core of indeterminacy in legal language.' In Gotti M, Bhatia V & Heller D (eds.) *Vagueness in normative texts*. New York: Peter Lang.

Walter H (1994). *L'aventure des langues en Occident*. Paris: Robert Laffont.

Woodbine N (1943). 'The language of English law.' *Speculum (18),* 400–420.

Le Roux, Pierre (1874–1975)

P-Y Lambert

This article is reproduced from the previous edition, volume 4, pp. 2061–2062 © 1994, Elsevier Ltd.

Born in 1874 in Plouëc (Côtes du Nord), Pierre Le Roux was a native speaker of the Breton language. He went to the University of Rennes, where he had Joseph Loth (*see* **Loth, Joseph (1847–1934)**) and Georges Dottin as his professors. His first work was a phonetic description of consonant mutations in his own dialect (Pleubian), a study published in 1896. With the help of a scholarship (Fondation Thiers), he came to Paris in 1898, and attended the lectures of the Celticists Gaidoz and D'Arbois de Jubainville. He succeeded Joseph Loth as teacher of Celtic in Rennes University in 1910.

For his thesis, he wrote a historical description of the Breton verb (1930), which qualified him as Professor. This book is an invaluable description of all the paradigms of the verb, from a historical and comparative point of view. It has never been replaced either in respect to the extensiveness of its sources or the accuracy of his analyses. For many of his interpretations (and particularly for the description of the verbs meaning 'to be'), Pierre Le Roux was inspired by two great precursors: Emile Ernault and François Vallée.

At the university, Pierre Le Roux started his dialectological investigations (1911–1913, resumed after the war in 1920–1921). He used to travel on his bicycle across Brittany, sometimes with cumbersome recording instruments. His results were published as a Linguistic Atlas of the Breton-speaking part of Brittany (the first fascicle was published in 1924, the 6th in 1963 with the help of Canon Falc'hun (*see* **Falc'hun, Francois (1909–1991)**). The questionnaire included more than 1200 questions, but only 680 maps have been published. He prepared the questionnaire in Rennes with the help of young Breton-speaking soldiers from different parts of Brittany. He had to limit himself to 78 places of inquiry, a rather loose network compared with the number of Breton-speaking communities (600 parishes).

This Linguistic Atlas recorded forms from dialects that are now extinct, for example, those of Bourg-de-Batz, Belle-Ile, and some places east of Vannes. The informant was chosen very carefully, with an investigation of the birthplace of parents and spouse. In some cases, however, Pierre Le Roux was obliged to rely on the hasty answers of people he was necessarily in touch with: tenants of inns and public houses. All in all, the documents gathered by Pierre Le Roux are an extraordinary witness to the Breton language spoken during the period. Even in the places where Breton is still spoken, it is certainly impoverished in its phonetics, lexicon, morphology, and syntax. The Linguistic Atlas inspired several new investigations into Breton dialectology and phonetics, particularly by François Falc'hun, who studied on this basis the linguistic geography of Brittany, and by Pierre Trépos, whose thesis on the morphology of the plural relied also on the data of the Linguistic Atlas.

Soon after becoming Dean of the Faculty of Arts at Rennes in 1944, Pierre Le Roux retired and lived quietly for the rest of his life; he died a centenarian in 1975. Among his publications were also some editions of older manuscripts (Breton songs of the Penguern collection).

See also: Breton; Celtic; Dialect Atlases; Falc'hun, Francois (1909–1991); Loth, Joseph (1847–1934).

Bibliography

Fleuriot L (1976–1977). 'Obituary of Pierre Le Roux.' *Etudes Celtiques 15(1),* 335.

Le Roux P (1924–1963). *Atlas Linguistique de la Basse-Bretagne,* fasc. 1, 1924; fasc. 2, 1927; fasc. 3, 1937; fasc. 4, 1943; fasc. 5, 1953; fasc. 6, 1963. Paris: Champion.

Le Roux P (1930). *Le verbe breton (morphologie, syntaxe).* Rennes. (2nd edn., 1957. Paris: Champion.)

Learners' Dictionaries

M Rundell, Lexicography MasterClass, Canterbury, UK

The monolingual learners' dictionary (MLD) – a type of dictionary designed specifically to meet the needs of non-native language learners – is above all an English phenomenon or, even more specifically, a British one. This is not to deny that there have been significant contributions in other languages and traditions. For example, Larousse's *Dictionnaire du français langue étrangère* (1987) was an interesting and in many ways innovative addition to the MLD canon. But it failed to establish a durable parallel thread to complement developments in English pedagogical lexicography. More recently, new learners' dictionaries have appeared for a number of European languages, including: Langenscheidt's *Großwörterbuch Deutsch als Fremdsprache* (1993), the Vox *Diccionario para la Enseñanza de la Lengua Española* (1997), Van Dale's *Pocketwoordenboek nederlands als tweede taal* (2003), and Nathan's *Dictionnaire de didactique du français langue étrangère et seconde* (2003). To a significant degree, however, these have tended to adopt the main features of the dominant British model. In the Anglophone world, meanwhile, several American publishers have entered the fray – with books such as *The American Heritage English as a Second Language Dictionary* (Houghton Mifflin, 1998) and *Random House Webster's dictionary of American English* (1997) – but their efforts have for the most part been highly conservative and have contributed little of interest to the field. On the theoretical front, our understanding of the lexicon has been greatly advanced by work in French, notably in the elaboration of lexicon grammars by Maurice Gross and his colleagues (*see* **Lexicon Grammars**). Nevertheless, the most powerful impetus in the development of MLDs has been – and remains – the practical challenge of providing language learners with the resources to meet their twin communicative needs: 'receptive' understanding and 'productive' use of a second language.

The foundations of the MLD were laid down in the 1920s and 1930s by a handful of British academics, including Harold Palmer, Michael West, and A. S. Hornby, whose theoretical interests in syntax, phraseology, and what would later be called collocation were motivated above all by their work in English language teaching. The story of the modern English learners' dictionary begins in 1935, with the publication of West's *New Method English Dictionary* (*NMED*), a fairly slim volume covering around 18 000 headwords. But it was Hornby's larger *Idiomatic and Syntactic English dictionary* of 1942 that was to prove the more enduring model – evolving as it did into the *Oxford Advanced Learner's Dictionary* (*OALD*, now in its seventh edition), which formed the template for most subsequent, major English dictionaries for advanced learners. Over 60 years on – despite a highly competitive market that has delivered an impressive series of innovations – Hornby's blueprint survives substantially intact. Numerous publishers have entered the field (and with over a billion people learning English worldwide, this is hardly surprising), and have experimented with a wide range of general and specialized learners' dictionaries for all levels of proficiency. The advanced-level MLD, however, remains the dominant format. Since the 1970s, Hornby's dictionary has been joined by several other U.K.-based contenders: the *Longman Dictionary of Contemporary English* (*LDOCE*, 1978), the *Collins COBUILD English Language Dictionary* (*COBUILD*, 1987), the *Cambridge International Dictionary of English* (*CIDE*, 1995), and the *Macmillan English Dictionary for Advanced Learners* (*MED*, 2002). Yet amid all this diversification, the trend is toward ever closer convergence in both form and content. Quite major upheavals (notably the publication by Collins of the first *COBUILD* dictionary in 1987) have been assimilated into an increasingly homogeneous model, as last year's innovation becomes this year's standard feature. Though experts will perceive numerous differences among the main players, what the average user sees is a set of very similar dictionaries that share a great many core features. To give a single (and fairly superficial) example of this convergent trend, the fourth edition of the large *COBUILD* dictionary (2003) has added the term 'advanced learner's dictionary' to its title (along with the MLDs produced by Oxford and Cambridge), has shrunk its (originally larger) format to 'standard MLD' size, and has, for the first time, included illustrations – thus bringing itself more closely into line with the other main dictionaries of this type.

Main Features of the MLD: A Summary

The features that have become a standard part of the MLD are summarized here, and the most important of these will be elaborated in the next section.

In terms of the dictionary's receptive function (helping users understand vocabulary items they have encountered):

- definitions are written in simple language, typically using a controlled 'defining vocabulary' of basic words;
- phraseology gets special attention: the MLD generally goes further than other types of dictionary in explaining the meanings not only of individual words but also of multiword expressions of every type;
- aids to navigation are provided (in the form of reduced definitions, appearing as 'signposts' or in menus) in order to minimize the known problems involved in locating the 'right' information.

In terms of the dictionary's productive function (helping users to make appropriate lexical choices and then use the chosen item 'correctly' and idiomatically):

- syntactic information provides users with a full account of the valence patterns that a given word or meaning enters into;
- example sentences are used extensively to show typical contexts of use and provide reliable models for production;
- considerable attention is paid to sociolinguistic features such as register, regional variety, and speaker attitude;
- usage notes provide supplementary guidance, e.g., facilitating lexical choices by disambiguating sets of close synonyms or warning against known sources of grammatical error.

In More Detail

Definitions

Here, three features distinguish the MLD from other dictionaries:

1. The use of a controlled 'defining vocabulary' (DV), enabling (perhaps a better word would be 'forcing') lexicographers to explain meanings through a limited list of basic, high-frequency words. From the beginning, a simplified defining language was one of the central features of the learners' dictionary. Three of the early pioneers in this field, Harold Palmer, A. S. Hornby, and Michael West, had a special interest in the notion of 'vocabulary control,' which arose – as Cowie (1999: 15) points out – "from a simple pedagogical need: to reduce the effort required to learn a foreign language by identifying those (relatively few) words which carried the main burden of communication." The process became more formalized when Longman – drawing on West (1953) – introduced a 2000-word DV in the first edition of the *Longman Dictionary of Contemporary English* (*LDOCE*) (1978). This approach was gradually emulated, so that at the time of writing four of the five major advanced MLDs employ a DV limited to 2000–3000 common words.
2. A preference for natural language: the past 15 years or so have seen a marked shift away from the formal, compressed, and often idiomatic language of traditional lexicography, toward a user-friendly style that conforms more closely to 'normal' prose (see esp. Rundell, 1998: 332–334). The following comparisons make the point:

NMED, 1935	**pain:** suffering of body or mind
*OALD*6, 2000	**pain: 1.** the feelings that you have in your body when you have been hurt or when you are ill **2.** mental or emotional suffering
*OALD*3, 1974	**aggressive:** quarrelsome; disposed to attack
*LDOCE*4, 2003	**aggressive:** behaving in an angry or threatening way, as if you want to attack someone

3. 'Full-sentence' defining: one of the major innovations in *COBUILD* (1987) was to abandon the 'substitutable' definition of traditional lexicography in favor of a defining sentence which includes the headword itself in its typical lexico-syntactic context. (The genesis and rationale of this approach is explained in Hanks, 1987, and discussed further in Sinclair, 2004.) To illustrate the difference this can make, a conventional definition might explain the phrasal verb *lay up* as 'to cause sb to stay in bed, not be able to work etc.' (*OALD*4, 1989). But in the *COBUILD* version ('If someone is **laid up** with an illness, the illness makes it necessary for them to stay in bed': *COBUILD*4, 2003), the structure of the definition matches the passive form in which this verb usually appears, while the prepositional phrase that typically accompanies it is fully integrated into the explanation. Though the *COBUILD* dictionaries are unique in employing this defining style in all cases, the full-sentence definition has an established place in the repertoire of definition types available to compilers in most of the English MLDs.

Syntax

Since Harold Palmer introduced his system of 'verb patterns' in 1938, descriptions of syntactic behavior have been a core feature of the MLD. Earlier MLDs were characterized by the use of elaborate coding schemas which, though descriptively powerful, were difficult to learn and mutually incompatible. Palmer's system was taken up and refined in *OALD*, while

LDOCE (1978) introduced an entirely different (but equally complex) approach based on alphanumeric codes. More recently, the emphasis has shifted toward a simpler, surface-grammar model which – while sacrificing some of the delicacy of earlier systems – assumes very little grammatical knowledge on the part of users. Compare the following accounts of the syntactic properties of the verb *recall*:

*OALD*3, 1974	[VP6A, C, 8,9,10, 19C] bring back to the mind; recollect: *I don't recall his name/face/meeting him/where I met him.*
LDOCE, 1978	[Wv6; T1,4,5,6a,b] to remember: *I can't recall his face/seeing him/that he came/where he lives/how to do it.*
MED, 2002	[I/T] to remember something: *Twenty years later, he could still recall the event.* ♦ +(**that**) *I seem to recall that you said you would do that yesterday.* ♦ +**who/where/why etc.** *Stephen frowned, trying to recall what had happened.* ♦ **recall doing sth** *I don't recall seeing the document.*

The economy of the older systems allows them to encode every *possible* pattern for a given meaning, regardless of its frequency: thus, the *OALD*3 code VP19C indicates the construction V+noun/pronoun/possessive+gerund (*I don't recall you saying that*), but – tellingly – the pattern is not exemplified. And in fact, corpus data shows it to be extremely rare, with just five occurrences in the 5500 or so instances of the verb *recall* on the British National Corpus (BNC). (Conversely, the simple V+gerund pattern – encoded as VP6C in *OALD*3 – appears over 150 times in the BNC.) The current approach – exemplified here by *MED* but equally likely to be found in other MLDs – emphasizes what is typical over what is possible, and uses corpus-derived examples to complement the user-friendly patterning information. And as with most other features of these dictionaries, there is now a fairly high degree of convergence among the various MLDs in their approach to describing syntax.

Examples

Dictionaries for native speakers (especially historical dictionaries) have generally used citations for purposes of illustration and attestation. By contrast, the examples in learners' dictionaries have always had a consciously pedagogical function. In the earlier MLDs, there was little emphasis on replicating actual performance: rather, examples tended toward the formulaic, often taking the form of short noun phrases or infinitive statements ('*inordinate desire*,' '*a football fixture*,' '*to instil knowledge*': all *NMED*, 1935), or merging several patterns into a single 'sentence' (see examples above, under 'Syntax'). The current range of MLDs, however, favor fuller and more natural examples taken from a corpus (or at least closely reflecting what the corpus shows). For a while (following the publication of *COBUILD* in 1987 and the advent of corpus data), there appeared to be a philosophical split between those who used 'made-up' examples and those who advocated the use of unmodified corpus lines (see e.g., Laufer, 1992; Potter, 1998). This debate had some mileage at a time when the use and availability of corpora was still limited, but it is no longer really relevant. (It also, very misleadingly, encouraged the perception that the main value of a corpus was as a source of dictionary examples.) Now that all serious MLDs use corpus data as their primary source of linguistic evidence, methodological differences among the various editorial units, in terms of how examples are chosen for the dictionary, have become fairly minimal.

The Impact of Corpora

Learners' dictionaries have introduced many innovations which have subsequently influenced other areas of lexicography. By far the most important of these, it goes without saying, is the systematic use of corpus data as the primary basis of dictionary text. Pioneered by John Sinclair and his colleagues in Birmingham, corpus lexicography burst upon the world in the form of the first *COBUILD* dictionary (1987), and the methodology was quickly taken up by other dictionary developers of all types. For those working at the coal face, the change in working practices has been almost total. As far as dictionary content goes, improvements in quality (above all in the area of semantic analysis) have been very significant. Corpus data has also enabled lexicographers to provide categories of information which were simply not available before. Notable examples are the explicit indicators of word frequency which now appear in most advanced MLDs (see e.g., Kilgarriff, 1997), and the lists of typical collocates shown first in *MED* (2002) and already becoming a standard feature (Kilgarriff and Rundell, 2002). These are enormous changes by any standards, and the longer-term impact of the corpus revolution (in terms of its challenges to existing linguistic assumptions and categories) is still far from clear. Yet despite all this, the MLD remains recognizably the same object as it was before the arrival of corpora.

Signs of Change

We are now, however, seeing signs that this durable paradigm is beginning to break down. New technology – supplying vastly improved linguistic data, and at the same opening up new media for delivering reference materials to users – has created conditions which threaten the well-established MLD model but at the same time present extraordinary opportunities for producing better dictionaries and for customizing them to more effectively meet the needs of users.

The advanced MLD has always been a global product: a general-purpose dictionary intended for learners from widely different linguistic and cultural backgrounds. It now seems unlikely that this format will survive indefinitely. Already we see some erosion of this position with the publication of country-specific editions, which offer anything from minor levels of customization – such as a Chinese or Japanese introduction with user guides in the 'local' language – to full-scale 'bilingualization,' with L1 equivalents supplied for each sense and translations of the example sentences. This trend looks set to intensify, given the possibilities offered by electronic media, and the global MLD might in future become just one among many types of learners' dictionary.

For most of its lifetime, the MLD appeared only in the form of a printed book, but learners' dictionaries are now routinely available on CD-ROM and handheld devices, and many also have an online presence. This is a rapidly changing scenario, and exactly how it will play out is anybody's guess. There will probably always be a demand – in some parts of the world at least – for printed dictionaries, but it is safe to assume that some form of electronic platform will become the primary delivery medium for the MLD in its most important markets. The possibilities here are only just beginning to be exploited. Most electronic dictionaries have replicated the content and structure of their printed originals, simply adding friendly interfaces and increasingly smart search functionality. But the next generation is now emerging – a good example being the CD-ROM version of *LDOCE*4 (2003), which gives users access to a 'hinterland' of additional examples and corpus citations for almost every word and phrase. Links between dictionary and corpus are just one of the possible ways in which the MLD will develop from here, and we may look back on the first decade of the new millennium as the point at which the one-size-fits-all MLD began to diversify into a range of more specialized reference resources.

See also: Computers in Lexicography; Corpus Lexicography; Definition; Hornby, Albert Sidney (1898–1978); Lexicography: Overview; Lexicon Grammars; Palmer, Harold Edward (1877–1949); Second and Foreign Language Learning and Teaching; Valency Grammar.

Bibliography

Dictionaries Cited

CIDE (1995). *Cambridge International Dictionary of English.* Procter P (ed.). Cambridge: Cambridge University Press.
COBUILD (1987). *Collins COBUILD English Language Dictionary.* Sinclair J (ed.). London/Glasgow: Collins.
*COBUILD*4 (2003). *Collins COBUILD Advanced Learner's English Dictionary.* Sinclair J (ed.). London/Glasgow: HarperCollins.
Hornby A S, Gatenby E V & Wakefield H (1942). *Idiomatic and Syntactic English Dictionary.* Tokyo: Kaitakusha.
LDOCE (1978). *Longman Dictionary of Contemporary English.* Procter P (ed.). London: Longman.
*LDOCE*4 (2003). *Longman Dictionary of Contemporary English* (4th edn.). Summers D (ed.). London: Longman.
MED (2002). *Macmillan English Dictionary for Advanced Learners.* Rundell M (ed.). Oxford: Macmillan.
NMED (1935). *The New Method English Dictionary,* by West M & Endicott J G. London: Longmans, Green.
*OALD*3 (1974). *Oxford Advanced Learner's Dictionary of Current English.* (3rd edn.), by Hornby A S, Cowie A P & Windsor Lewis J. Oxford: Oxford University Press.
*OALD*4 (1989). *Oxford Advanced Learner's Dictionary of Current English* (4th edn.). Cowie A P (ed.). Oxford: Oxford University Press.
*OALD*6 (2000). *Oxford Advanced Learner's Dictionary of Current English* (6th edn.). Wehmeier S (ed.). Oxford: Oxford University Press.
Palmer H E (1938). *A Grammar of English Words.* London: Longmans, Green.

Other References

Cowie A P (1999). *English Dictionaries for Foreign Learners.* Oxford: Oxford University Press.
Hanks P (1987). 'Definitions and explanations.' In Sinclair J M (ed.) *Looking Up.* London: Collins. 116–136.
Harras G, Hass U & Strauss G (1991). *Wortbedeutung und ihre Darstellung im Wörterbücher.* Berlin/New York: De Gruyter.
Kilgarriff A (1997). 'Putting frequencies in the dictionary.' *International Journal of Lexicography 10(2),* 135–155.
Kilgarriff A & Rundell M (2002). 'Lexical profiling software and its lexicographic applications: a case study.' In Braasch A & Povlsen C (eds.) *Proceedings of the Tenth EURALEX Congress.* Copenhagen: Center for Sprogteknologi. 807–818.
Laufer B (1992). 'The effect of dictionary definitions and examples on the use and comprehension of new L2 words.' *Cahiers de Lexicologie 63,* 131–142.

Potter L (1998). 'Setting a good example.' In Fontenelle T, Hiligsmann P, Michiels A, Moulin A & Theissen S (eds.) *EURALEX '98 Proceedings*. Liège: University of Liège. 357–362.
Rundell M (1998). 'Recent trends in English pedagogical lexicography.' *International Journal of Lexicography 11(4)*, 315–342.
Sinclair J (2004). 'In praise of the dictionary.' In Williams G & Vessier S (eds.) *Proceedings of the Eleventh EURALEX Congress*. Lorient, France: Université de Bretagne Sud. 1–11.
West M (1953). *A General Service List of English Words*. London: Longmans, Green.

Learning Sign Language as a Second Language

R I Mayberry, University of California, San Diego, CA, USA

Sign languages are natural languages. This means that sign languages are often learned as second languages, just like spoken languages. Unlike spoken languages, however, sign languages are expressed with the hands, arms, and face and understood through the eyes. This means there are two kinds of second sign language learning: unimodal second language learning, where a person knows two or more sign languages, and bimodal second language learning, where a person knows one or more sign languages and one or more spoken languages. An important question, then, is whether the sensory-modality of the first and second languages changes what we know about second language learning over the life span. Another important question is whether and how the age of first language learning affects the ultimate outcome of second language learning. Clearly several factors in the second language must be considered in any account of second language learning of sign languages, sensorimotor modality, the temporal order of the first and second language learning, and the ages when the first and second languages are learned. Here we focus first on second language (L2) learning when the first language (L1) learning begins in early childhood. We consider both simultaneous and sequential L1 and L2 learning. Next we consider L2 learning in the case where L1 learning begins after early childhood.

Learning Two Languages in Infancy

Babies whose parents use a sign language with them in infancy often learn two languages simultaneously. One kind of L2 learning is by hearing babies whose deaf parents use a sign language with them. They typically acquire both a sign language and a spoken language. This is an example of bimodal bilingualism. The ages when these babies reach their early language milestones – such as the appearance of the first word, the first 50 words, and the first word combinations – are similar to those of babies learning two spoken languages and babies learning only one language, signed or spoken (Petitto *et al.*, 2001). The early vocabularies of babies acquiring both a signed and a spoken language are semantically similar across the two languages, although the specific vocabulary items are not identical. Fewer than a third of the vocabulary items the child knows in each language are translation equivalents, similar to children acquiring two spoken languages. The vocabulary sizes of babies learning a signed and a spoken language simultaneously are similar to those of babies learning only one spoken language, especially when the lexical items of the two languages are considered together, as is the case for babies learning two spoken languages (Holowka *et al.*, 2002; Pearson *et al.*, 1993).

Babies who are deaf and whose parents use a sign language with them from birth often learn an L2 at the same time. The L2 may be another sign language, but unfortunately little is known about bilingualism in the visual modality. More is known about bimodal bilingualism. For example, when learning both SLN (Sign Language of the Netherlands) and Dutch, one deaf baby was observed to begin learning vocabulary in SLN. The child then used the sign vocabulary to help learn Dutch vocabulary items up to 36 months. In an intermediate stage, the child both signed and spoke vocabulary items. The child eventually separated the vocabularies of the two languages and either expressed signs without speaking or spoke words without signing. Around the same age, the child showed an awareness that the word order patterns of SLN and Dutch are not the same (Hoiting, 1998).

Learning Two Languages Sequentially

Adults and adolescents frequently learn sign languages as second languages. Hearing adults may learn sign languages for professional reasons. Teachers, interpreters, therapists, and researchers working with

Potter L (1998). 'Setting a good example.' In Fontenelle T, Hiligsmann P, Michiels A, Moulin A & Theissen S (eds.) *EURALEX '98 Proceedings*. Liège: University of Liège. 357–362.

Rundell M (1998). 'Recent trends in English pedagogical lexicography.' *International Journal of Lexicography 11(4)*, 315–342.

Sinclair J (2004). 'In praise of the dictionary.' In Williams G & Vessier S (eds.) *Proceedings of the Eleventh EURALEX Congress*. Lorient, France: Université de Bretagne Sud. 1–11.

West M (1953). *A General Service List of English Words*. London: Longmans, Green.

Learning Sign Language as a Second Language

R I Mayberry, University of California, San Diego, CA, USA

Sign languages are natural languages. This means that sign languages are often learned as second languages, just like spoken languages. Unlike spoken languages, however, sign languages are expressed with the hands, arms, and face and understood through the eyes. This means there are two kinds of second sign language learning: unimodal second language learning, where a person knows two or more sign languages, and bimodal second language learning, where a person knows one or more sign languages and one or more spoken languages. An important question, then, is whether the sensory-modality of the first and second languages changes what we know about second language learning over the life span. Another important question is whether and how the age of first language learning affects the ultimate outcome of second language learning. Clearly several factors in the second language must be considered in any account of second language learning of sign languages, sensorimotor modality, the temporal order of the first and second language learning, and the ages when the first and second languages are learned. Here we focus first on second language (L2) learning when the first language (L1) learning begins in early childhood. We consider both simultaneous and sequential L1 and L2 learning. Next we consider L2 learning in the case where L1 learning begins after early childhood.

Learning Two Languages in Infancy

Babies whose parents use a sign language with them in infancy often learn two languages simultaneously. One kind of L2 learning is by hearing babies whose deaf parents use a sign language with them. They typically acquire both a sign language and a spoken language. This is an example of bimodal bilingualism. The ages when these babies reach their early language milestones – such as the appearance of the first word, the first 50 words, and the first word combinations – are similar to those of babies learning two spoken languages and babies learning only one language, signed or spoken (Petitto *et al.*, 2001). The early vocabularies of babies acquiring both a signed and a spoken language are semantically similar across the two languages, although the specific vocabulary items are not identical. Fewer than a third of the vocabulary items the child knows in each language are translation equivalents, similar to children acquiring two spoken languages. The vocabulary sizes of babies learning a signed and a spoken language simultaneously are similar to those of babies learning only one spoken language, especially when the lexical items of the two languages are considered together, as is the case for babies learning two spoken languages (Holowka *et al.*, 2002; Pearson *et al.*, 1993).

Babies who are deaf and whose parents use a sign language with them from birth often learn an L2 at the same time. The L2 may be another sign language, but unfortunately little is known about bilingualism in the visual modality. More is known about bimodal bilingualism. For example, when learning both SLN (Sign Language of the Netherlands) and Dutch, one deaf baby was observed to begin learning vocabulary in SLN. The child then used the sign vocabulary to help learn Dutch vocabulary items up to 36 months. In an intermediate stage, the child both signed and spoke vocabulary items. The child eventually separated the vocabularies of the two languages and either expressed signs without speaking or spoke words without signing. Around the same age, the child showed an awareness that the word order patterns of SLN and Dutch are not the same (Hoiting, 1998).

Learning Two Languages Sequentially

Adults and adolescents frequently learn sign languages as second languages. Hearing adults may learn sign languages for professional reasons. Teachers, interpreters, therapists, and researchers working with

children and adults who are deaf often learn sign language in university courses designed for them. The first stage of adult L2 learning of sign languages is characterized by a marked reliance on the perceived iconic features of signs. Beginning signers use iconicity as a mnemonic aid to remember new vocabulary items for both sign expression and recognition (Campbell *et al.*, 1992). Beginning sign learners often want to know the mimetic 'reason' for every sign form. Although sign languages are rooted in manual gesture, they are not iconic by nature. To the contrary, all sign languages have phonological structure, that is, meaningless articulatory units that combine to make the words of the language. Research shows that signers manipulate the phonological structure of sign languages during language comprehension and expression, just as phonological structure is a basic part of listening to and speaking a spoken language.

A key difference between adult L2 learning and infant learning of sign languages is that iconicity plays no role in infants' vocabulary learning. Moreover, iconicity plays no role in the human mind's comprehension of sign language for both L2 and L1 signers who are highly proficient. Beginning signers may recruit what they already know that is close in nature to sign language, namely in gesture and pantomime. This may help beginning signers learn how to position their fingers, hands, and arms in making signs and serve as a means to remember the meanings of the signs they are seeing for the first time. The strategy may be analogous to beginning speakers of a second language using the sound patterns of their native language to help themselves remember the sound patterns of a new spoken L2 they are learning. Clearly L2 learners abandon this strategy once they learn the phonological structure of the sign language.

Beyond vocabulary learning, L2 learners of sign languages must also master morphology and syntax. One problem for L2 learners is that the syntax and morphology of natural sign languages are not always taught in sign language courses. Many courses primarily teach the frozen lexicon of sign languages, excluding classifier constructions, along with finger spelling. This teaching method inadvertently encourages beginning L2 learners to express sign language vocabulary using the word order patterns of their spoken L1. The result is a kind of pidgin where L2 content words (vocabulary for people, actions, internal states, and objects) are signed according to the word order of the L1 and devoid of the morphology and syntax of any language.

Adults and adolescents who are deaf and who acquired an L1 in early childhood often learn sign languages as second languages. For example, L1 learners of a sign language may later learn some other sign language as an L2, an example of bilingualism in the visual modality. This is common where sign languages are in close geographical proximity, as is the case for LSQ (Langue des signes québeçoise) and ASL (American Sign Language) in Canada. Another possible example of bilingualism in the visual modality is the L2 sign language learning by deaf adults and adolescents who are proficient in a spoken L1, which is often in a read and lip-read form. These L2 learners are more likely to immerse themselves in Deaf communities than hearing L2 learners of sign languages and for this reason are more likely to achieve higher levels of ultimate proficiency than their hearing L2 peers.

Another type of sequential L2 learning of sign languages is found in educational projects that teach BSL (British Sign Language) or ASL (American Sign Language) to kindergarten and elementary school children as a second language (Daniels, 2003). Other programs teach sign languages to hearing elementary school children to enhance their reading and spatial cognitive skills. Yet other programs use signs with hearing children who are delayed in their spoken language development, such as children with autism or Down syndrome. The sign learning in the latter instances is more properly considered L1 rather than L2 learning, however.

Delayed L1 Learning and L2 Learning

The common definition of L2 learning is the learning of a second language simultaneously with, or subsequent to, learning a first language. Although L1 learning in early childhood is the norm for babies born with normal hearing, because they are immersed in spoken language, it is not the norm for babies born deaf. A minority learn sign language from birth because their parents use a sign language with them, as discussed earlier in this article. However, the majority of babies born deaf are isolated not only from the languages spoken around them but also from sign languages if their parents do not sign with them. Even after the initial diagnosis of hearing loss and onset of intervention services for the deaf child, sign languages are often withheld from them in the belief that doing so encourages deaf babies to learn to speak. Note that this is contrary to the popular trend in North America of teaching signs to hearing toddlers because it reduces frustration for the parent and child (Goodwyn *et al.*, 2000), and to the clinical case reports of deaf babies making faster progress learning spoken words after they have learned a sign lexicon. The important point to remember is that sign language learning for many deaf children and

adolescents is not L2 learning but is in reality L1 learning begun at an abnormally late age, i.e., after early childhood.

Learning an L1 for the first time after early childhood has a plethora of negative effects on adult language comprehension. For example, the ability to comprehend and remember sign language sentences and stories declines as the age of L1 sign language learning increases. This is apparent in the kinds of lexical mistakes that delayed L1 learners make when engaged in sign language comprehension and production tasks. They often make phonological errors, producing lexical items that are real signs but with forms that violate the semantic and syntactic framework of the sentence they are reproducing (see Figure 1). By contrast, L1 learners who learned sign language from birth, or shortly thereafter, make few lexical errors, and the ones they do make maintain the syntactic and semantic frame of the sentence they are reproducing. Delayed L1 learners also show a reduced ability to comprehend sign language.

The effects of delayed L1 learning are also apparent on grammatical knowledge of sign language. For example, delayed L1 learners are more likely to delete grammatical inflections from signs when producing sign language sentences, whereas early L1 learners either inflect signs or express the same meanings in separate lexical units. Highly delayed L1 learners have been observed to perform at near chance levels on grammatical tasks in sign language.

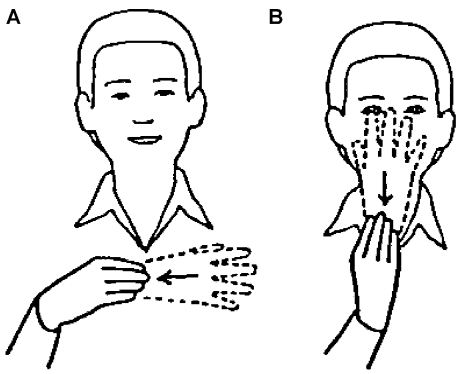


Figure 1 Panel A shows the sign AND in the ASL sentence 'I ate too much turkey AND potato,' which was the target sentence. Panel B shows the phonologically based substitution error, which resulted in the sentence 'I ate too much turkey SLEEP potato.' Note that the target and error differ in only one phonological parameter: place of articulation. (From Mayberry R 1996. 'The importance of childhood to language acquisition.' In Goodman J C & Nusbaum J C (eds.) *The development of speech perception*, p. 68. Cambridge: MIT Press. Illustration by Betty Raskin © R. Mayberry, with permission.)

By contrast, early L1 learners of sign languages tend to be grammatically accurate.

Delayed L1 learning not only affects the ultimate proficiency with which the L1 can be understood and expressed but also affects the ultimate outcome of L2 learning crossmodally. Early L1 learners of sign languages often show near-native levels of L2 proficiency. For example, people who were born deaf and acquired a sign language early in life often show high levels of L2 learning of spoken languages in read and lip-read forms. By contrast, people who were born deaf and did not experience a fully perceptible language until after early childhood, i.e., delayed L1 learners, typically show low levels of L2 spoken language proficiency in read and lip-read forms. In other words, there is a strong relationship between the age when the L1 is learned and the L2 outcome, independent of the sensorimotor modalities of the languages. Although the linguistic, cognitive, or neural reasons that early L1 learning is necessary for L2 learning to be successful are not understood at present, it is clear that any discussion of L2 learning of sign languages requires consideration of the age when the L1 learning begins.

We return now to the two questions posed at the beginning of this article. First, does the sensorimotor modality of the first and second languages change what we know about how second languages are learned over the life span? The second question is how the age of first language learning affects the ultimate outcome of second language learning. Much of what we know about second language learning of sign languages suggests that there are few differences between second language learning of sign languages and of spoken languages, as has been previously found for infant L1 learning of sign languages compared with spoken languages. One exception appears to be the initial iconic strategy that adult second language learners of sign languages use to begin to parse and remember the phonological forms of signs. The second crucial difference is the age when individuals born deaf first experience a fully perceptible language. It appears that the timing of first language learning during the life span sets the course for all subsequent language learning regardless of the sensorimotor modality of the language learning.

See also: Bilingual Language Development: Early Years; Disorders of Sign Language; Iconicity: Sign Language; Language Development in Deaf Children with Hearing Parents; Language Development: Overview; Language Education of the Deaf; Modality Issues in Signed and Spoken Language; Sign Language: Acquisition; Sign Language: Interpreting; Sign Language: Overview.

Bibliography

Bartman S & Freeman N (2003). 'Teaching language to a two-year-old with autism.' *Journal on Developmental Disabilities 10(1),* 47–53.

Boudreault P & Mayberry R I (in press). 'Grammatical processing in American Sign Language: age of first-language acquisition effects in relation to syntactic structure.' *Language and Cognitive Processes.*

Campbell R, Martin P & White T (1992). 'Forced choice recognition of sign in novice learners of British Sign Language.' *Applied Linguistics 13(2),* 185–201.

Chamberlain C & Mayberry R I (2000). 'Theorizing about the relationship between ASL and reading.' In Chamberlain C, Morford J & Mayberry R I (eds.) *Language acquisition by eye.* Mahwah, NJ: Lawrence Erlbaum and Associates. 221–260.

Clibbens J (2001). 'Signing and lexical development in children with Down syndrome.' *Down Syndrome: Research and Practice 7(3),* 101–105.

Cochran B P, McDonald J L & Parault S J (1999). 'Too smart for their own good: the disadvantage of a superior processing capacity for adult language learners.' *Journal of Memory and Language 41,* 30–58.

Daniels M (2003). 'Using a signed language as a second language for kindergarten students.' *Child Study Journal 33(1),* 53069.

Goodwyn S W, Acredolo L P & Brown C A (2000). 'Impact of symbolic gesturing on early language development.' *Journal of Nonverbal Behavior 24(2),* 81–103.

Hoemann H W & Keske C M (1995). 'Proactive interference and language change in hearing adult students of American Sign Language.' *Sign Language Studies 86,* 45–62.

Hoemann H W & Koenig T J (1984). 'Categorical coding of manual and English alphabet characters by beginning students of American Sign Language.' *Sign Language Studies 44,* 175–181.

Hoiting N (1998). 'Early bilingual development in Sign Language of the Netherlands and Dutch: a case study of a deaf child.' In Clark E V (ed.) *The Proceedings of the 29th Annual Child Language Research Forum.* Chicago: Center for the Study of Language and Information. 73–80.

Holowka S, Brosseau-Lapre F & Petitto L A (2002). 'Semantic and conceptual knowledge underlying bilingual babies' first signs and words.' *Language Learning 52,* 205–262.

Lederberg A R & Spencer P E (in press). 'Critical periods in the acquisition of lexical skills: evidence from deaf individuals.' In Fletcher P & Miller J (eds.) *Language disorders and developmental theory.* Philadelphia: John Benjamins.

Mayberry R I (1993). 'First-language acquisition after childhood differs from second-language acquisition: the case of American Sign Language.' *Journal of Speech and Hearing Research 36,* 1258–1270.

Mayberry R I (2002). 'Cognitive development of deaf children: the interface of language and perception in neuropsychology.' In Segaolwitz S J & Rapin I (eds.) *Handbook of neuropsychology.* vol. 8, part 2 (2nd edn.). Amsterdam: Elsevier. 71–107.

Mayberry R I & Eichen E B (1991). 'The long-lasting advantage of learning sign language in childhood: another look at the critical period for language acquisition.' *Journal of Memory and Language 30,* 486–512.

Mayberry R I & Lock E (2003). 'Age constraints on first versus second language acquisition: evidence for linguistic plasticity and epigenesis.' *Brain and Language 87,* 369–383.

Mayberry R I, Lock E & Kazmi H (2002). 'Linguistic ability and early language exposure.' *Nature 417,* 38.

Newport E (1990). 'Maturational constraints on language learning.' *Cognitive Science 14,* 11–28.

Padden C & Ramsey C (2000). 'American Sign Language and reading ability.' In Chamberlain C, Morford J & Mayberry R I (eds.) *Language acquisition by eye.* Mahwah, NJ: Lawrence Erlbaum and Associates. 165–190.

Pearson B Z, Fernandez S & Oller D K (1993). 'Assessing lexical development in bilingual infants and toddlers: comparison to monolingual norms.' *Language Learning 43(1),* 93–120.

Pettito L A, Katerelos M, Levy B G, Gauna K, Tetreault K & Ferraro V (2001). 'Bilingual signed and spoken language acquisition from birth: implications for the mechanisms underlying early bilingual language acquisition.' *Journal of Child Language 28(2),* 453–496.

Schiff-Myers N (1988). 'Hearing children of deaf parents.' In Bishop D & Mogford K (eds.) *Language development in exceptional circumstances.* Edinburgh: Churchill Livingstone. 47–61.

Seaman C (1999). 'Language and culture choices of deaf adults.' *CAEDHH/ACESM 25(1),* 52–64.

Stedt J D, Salvia J & Nelson K E (1987). 'Using mnemonic explanations to improve sign lexicon learning with children and young adults.' *Sign Language Studies 55,* 141–161.

Strong M & Prinz P (1997). 'A study of the relationship between American Sign Language and English literacy.' *Journal of Deaf Studies and Deaf Education 2,* 37–46.

Yoshinaga-Itano C, Sedey A L, Coulter D K & Mehl A L (1998). 'Language of early- and late-identified children with hearing loss.' *American Academy of Pediatrics 102(5),* 1161–1171.